

**Universitat de Lleida**  
**Escola Tècnica Superior d'Enginyeria Agrària**  
**Proyecto Final de Máster**

**DIFERENTES SISTEMAS DE GENOTIPADO PARA  
DIFERENTES APLICACIONES EN LA MEJORA GENÉTICA DEL  
MELOCOTONERO**

**Alumno: Eduard Salvadó Peñalver**

**Tutor: Pilar Muñoz Odina (UdL)**

**Cotutor: Iban Eduardo Muñoz (CRAG)**

**Setiembre 2018**

## *Agradecimientos*

Me gustaría agradecer a **Pere Arús** que me diera la oportunidad de entrar en su grupo de *Rosaceae Genetics and Genomics*, en el Centre de Recerca Agrigenòmica (CRAG), para poder aprender un poco del apasionante mundo de la genética. A **Iban Eduardo Muñoz** le agradezco su gran dosis de paciencia, humor y su capacidad de transmitirme los conocimientos necesarios para desarrollar este trabajo. A **Pilar Muñoz Odena** le agradezco que me enseñará lo interesante que puede ser la biotecnología. Y por último he de mencionar a **Jordi García Mas** por ayudarme a encontrar un grupo de investigación para realizar este TFM.

Toda la parte práctica de este trabajo ha sido desarrollado en el Centre de Recerca Agrigenòmica (CRAG) en el programa de investigación de genómica de plantas y animales, en el grupo de *Rosaceae Genetics and Genomics* .

## RESUMEN

La mejora genética clásica en melocotonero es una metodología lenta y costosa. Por este motivo han aparecido nuevas metodologías basadas en marcadores moleculares, como la selección asistida por marcadores, que permiten aumentar su eficiencia. En este trabajo se presenta la prueba de concepto de una nueva estrategia basada en marcadores llamada Resíntesis. El objetivo es obtener individuos con la misma configuración genética que una variedad elite, pero con pequeños cambios que aporten nuevas características. Para ello se han genotipado 96 SNPs, 2 SSRs y el marcador InDel SRcod en 418 individuos de una población  $F_2$  de la variedad SweetDream (SD). Primero se ha construido un mapa genético para verificar la cobertura del genoma con los marcadores utilizados. Usando los genotipos obtenidos, se han seleccionado 6 individuos con un alto porcentaje del genoma con la misma configuración genética que el parental SD (Resíntesis A) y 4 parejas de individuos complementarios de forma que al cruzarlos se puedan obtener fácilmente individuos con una configuración genética similar a SD. El marcador SRcod nos ha permitido predecir qué individuos presentaran el carácter a descartar Slow Ripening (SR), que es una alteración en el proceso de maduración del fruto. En un futuro los individuos seleccionados (Res A) o sus descendientes (Res B), serán fenotipados para verificar si el individuo obtenido presenta alguna característica interesante, validando de esta forma la utilidad de la estrategia de la Resíntesis.

**Palabras clave:** Resíntesis, Marcadores moleculares, *Slow Ripening*.

## ABSTRACT

The classic genetic improvement in peach trees is a slow and expensive methodology. For this reason, new methodologies based on molecular markers, such as marker-assisted selection, have emerged to increase their efficiency. In this paper, we present the proof of concept of a new strategy based on markers called Resynthesis. The objective is to obtain individuals with the same genetic configuration as an elite variety, with small changes that provide new characteristics. To do this, 96 SNPs, 2 SSRs and the InDel SRcod marker were genotyped in 418 individuals of a  $F_2$  population of the SweetDream (SD) variety. First, a genetic map has been constructed to verify the coverage of the genome with the markers used. Using the genotypes obtained, 6 individuals have been selected with high percentage of the genome with the same genetic configuration as the SD parent (Resynthesis A) and 4 pairs of

complementary individuals so that when crossing them, individuals with a similar genetic SD configuration can be easily obtained. The SRcod marker has allowed us to predict which individuals presented the character Slow Ripening to discard, which is an alteration in the process of fruit ripening. In the future, the selected individuals (Res A) or their descendants (Res B), will be phenotyped if the individual obtained presents an interesting characteristic, validating in this way, the usefulness of the Resynthesis strategy.

**Key words:** Resynthesis, Molecular markers, Slow Ripening.

## RESUM

La millora genètica clàssica en presseguer es una metodologia lenta y costosa. Por aquest motiu han aparegut noves metodologies basades en marcadors moleculars, como la selecció assistida por marcadores, que permeten incrementar la seva eficiència. En este treball es presenta la prova de concepte de una nova estratègia basada en marcadors anomenada Resíntesis. El objectiu es obtenir individus amb la mateixa configuració genètica que una varietat elit, però amb petits canvis que proporcionin noves característiques. Per això, s'han genotipat 96 SNPs, 2 SSRs i el marcador InDel SRcod en 418 individus de una població F<sub>2</sub> de la varietat SweetDream (SD). Primer, s'ha construït un mapa genètic para verificar la cobertura del genoma amb els marcadors utilitzats. Fent servir el genotips obtinguts, s'han seleccionat 6 individus amb un alt percentatge del genoma amb la mateixa configuració genètica que el parental SD (Resíntesis A) y 4 parelles de individus complementaris de forma que al creuar-los es puguin obtenir fàcilment individus amb una configuració genètica similar a SD. El marcador SRcod ens ha permès preveure quins individus presentaran el caràcter a descartar Slow Ripening (SR), que es una alteració en el procés de maduració del fruit. En un futur, els individus seleccionats (Res A) o els seus descendents (Res B), seran fenotipats per a verificar si el individu obtingut presenta alguna característica interesant, validant d'aquesta forma la utilitat de la estratègia de la Resíntesis.

**Palabras clave:** Resíntesis, Marcadores moleculars, *Slow Ripening*.

# ÍNDICE

1	INTRODUCCIÓN.....	1
1.1	Mejora genética del melocotón .....	2
1.2	Selección asistida por marcadores moleculares .....	2
1.2.1	Marcadores moleculares .....	3
1.2.2	Mapas de ligamiento genético.....	6
1.2.3	Ejemplo de selección asistida por marcadores: maduración lenta o <i>slow ripening</i> (SR) 7	
1.2.4	Nuevas estrategias de mejora basadas en la selección asistida por marcadores: La Resíntesis .....	7
2	OBJETIVOS.....	9
3	MATERIALES Y MÉTODOS .....	10
3.1	Material vegetal.....	10
3.2	Extracción de ADN .....	10
3.3	Genotipado mediante SNPs (Single Nucleotide Polymorphism).....	11
3.4	Genotipado mediante microsatélites (SSR: <i>Simple Sequence Repeats</i> ).....	12
3.5	Genotipado del marcador SRcod.....	13
3.6	Tratamiento previo de los datos antes de realizar el mapa genético y la Resíntesis	13
3.7	Construcción de un mapa de ligamiento genético .....	14
3.8	Resíntesis .....	15
3.8.1	Metodología de estudio de la Resíntesis A.....	15
3.8.2	Metodología de estudio de la Resíntesis B .....	15
3.8.3	Método para encontrar los individuos aptos para realizar cruzamientos .	16
4	RESULTADOS .....	18
4.1	Material vegetal disponible para los ensayos .....	18

4.2	Genotipado .....	18
4.2.1	Genotipado mediante SNP .....	18
4.2.2	Genotipado mediante SSR .....	24
4.2.3	Genotipado slow ripening (SR) .....	26
4.2.4	Preselección de individuos.....	28
4.3	Mapa ligamiento genético.....	29
4.4	Individuos aptos para la Resíntesis A .....	31
4.4.1	Individuos aptos para la Resíntesis B.....	33
5	DISCUSIÓN .....	36
5.1	Genotipado mediante SNP .....	36
5.1.1	Perdida de datos ( <i>missing values</i> ) .....	36
5.2	Genotipado mediante SSR.....	37
5.3	Genotipado slow ripening (SR) .....	37
5.4	Mapa de ligamiento genético.....	38
5.5	Resíntesis .....	38
5.5.1	Resíntesis A .....	38
5.5.2	Resíntesis B .....	40
6	IMPACTO ECONÓMICO DE LOS RESULTADOS .....	44
6.1	Reducción de los tiempos y costes de selección varietal mediante el uso de marcadores moleculares.....	44
6.1.1	Caso de estudio de este trabajo .....	44
6.1.2	Ejemplo real de un estudio de mejora genética .....	45
7	CONCLUSIONES .....	49
8	BIBLIOGRAFÍA.....	50

# 1 INTRODUCCIÓN

El melocotón [*Prunus persica* L. (Batsch)] es una especie de árbol originario de China de la familia *Rosaceae*, que se cultiva desde hace unos 3.000 años. El cultivo del melocotón no tuvo mucha importancia en España hasta bien entrado el siglo XIX. Actualmente China es la primera productora de melocotones del mundo, seguido de Italia, Estados Unidos y de España que es la cuarta (FAOSTAT con datos del 2016)<sup>1</sup>, con una producción de 354.529 tn en el año 2017, según los datos de Cooperativas agroalimentarias de España (2018).

En España se cultiva principalmente en las CC.AA de Andalucía, Aragón, Catalunya, Extremadura y Valencia (Figura 1), debido básicamente a las exigencias climáticas del cultivo.

El melocotón requiere primaveras sin heladas, ya que el frío puede destruir las flores o frutos cuajados. El cultivo se adapta mejor en regiones templadas y en zonas donde no alcancen temperaturas extremadamente bajas en invierno, a -15°C se producen daños en las yemas de flor (Infoagro, 2002). Requiere de muchas horas de luz y no tolera bien los climas demasiado lluviosos porque es sensible a la asfixia radicular, con lo que terrenos con mucha fracción arcillosa y algo deprimidos, donde se pueda dar acumulación de agua, no serán los más aptos para desarrollar este cultivo.

Las variedades plantadas se escogen en función de la zona climática y la cantidad de horas-frío necesarias, se requieren de 400 a 800 horas y depende de la variedad. Para zonas con menos cantidad de horas-frío se escogen variedades extratempranas y en las de mayor cantidad de horas-frío se utilizan variedades tardías.

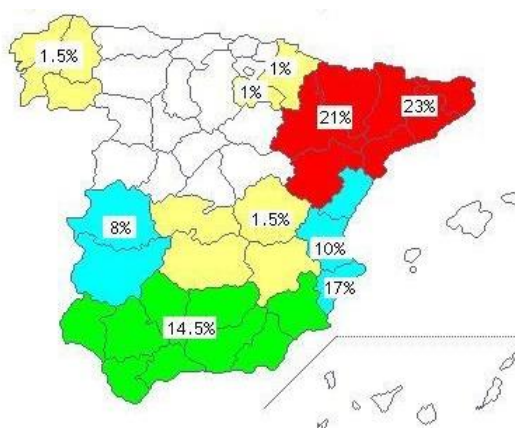


Figura 1 – Aportación porcentual de las Comunidades Autónomas a la superficie total cultivada en España del melocotonero. El mapa se ha elaborado con los datos del Anuario de estadística agroalimentaria del MARM 2015. Fuente: <https://www.mapama.gob.es/app/MaterialVegetal/fichaMaterialVegetal.aspx?idFicha=959>

<sup>1</sup> <http://www.fao.org/faostat/en/#data/QC/visualize>

## **1.1 Mejora genética del melocotón**

La mejora genética clásica ha conseguido que el cultivo del melocotón se expanda por diferentes regiones y se adapte a climas distintos, a pesar de que el melocotón presenta una escasa base genética, producto de su comportamiento autógamo (Aranzana MJ, 2010).

El método de mejora clásica para el cultivo del melocotón consiste básicamente en: recolectar el polen, realizar hibridaciones en campo, esperar a que el árbol de frutos y recolectarlos para cultivar embriones y/o germinar semillas, para posteriormente ser trasplantadas en campo y seleccionar los individuos en base a su fenotipo; con el coste en recursos y tiempo que conlleva evaluar cada árbol, ya que por lo menos se mantienen durante 4 años en el campo experimental.

Según Pérez (2012) el problema de la mejora clásica, es que existen dos factores limitantes: ciclos reproductivos largos y largos períodos de juvenilidad, con lo que registrar una variedad, mediante estas técnicas, puede conllevar 10 años de media. Para minimizar estos inconvenientes y aumentar la eficiencia del proceso de mejora en algunos casos se está recurriendo a la utilización de marcadores moleculares.

Actualmente las líneas de mejora van más allá de buscar mejores cualidades organolépticas para su uso comercial y se busca introducir resistencias a plagas y enfermedades, con lo que se vería reducido el uso de tratamientos fitosanitarios. Una de las mejoras más interesantes es poder seleccionar individuos en función del período de floración y maduración.

## **1.2 Selección asistida por marcadores moleculares**

Como alternativa a la selección por fenotipo tradicional existe la selección asistida por marcadores (SAM), que consiste en identificar una secuencia de ADN que está próxima al gen (o locus), donde tenemos algún rasgo de interés que queremos detectar. Utilizando esta secuencia podremos desarrollar un marcador molecular que permitirá seleccionar en base al genotipo en fases tempranas de la planta (primera hoja), lo que permite ahorrar recursos como superficie de cultivo, tratamientos y mano de obra (poda, fenotipado...).



### 1.2.1 Marcadores moleculares

Los marcadores moleculares corresponden a cualquier variación en la secuencia de ADN que puede ser detectable. Este tipo de marcadores pueden evaluarse desde que los individuos están en sus primeros estadios de desarrollo. (Solís & Andrade, 2015).

En resumen, cualquier característica A que esté asociada a la presencia o la expresión de otra característica B, se puede considerar un marcador, ya que la presencia de A implica la de B necesariamente.

Actualmente, por ejemplo, se han identificado marcadores para muchos caracteres como la forma del fruto (redondo/plano), el tipo de fruto (melocotón/nectarina), el color de la piel y de la pulpa, etc. Marcadores moleculares basados en SSR y SNPs

A finales de la década de los 80, tenemos los marcadores denominados microsatélites SSR<sup>2</sup> o STR (*Simple Sequence Repeats* o *Short Tandem Repeats*) y que vieron la luz por primera vez en los trabajos sobre el aislamiento y caracterización de microsatélites (Tautz, 1989).

Las secuencias de tipo microsatélite se componen por unidades cortas de 1 a 6 pares de bases, llamados motivos básicos, que se repiten un determinado número de veces (Figura 2). Cada secuencia se identifica por el tipo de unidad repetida y por el sitio que ocupan en el genoma (locus).

Para detectar las SSR utilizamos la amplificación por PCR<sup>3</sup> de la región que los contiene, empleando cebadores de las regiones flanqueantes. Los productos de la PCR se denominan amplicones y para poder visualizarlos se suele recurrir a la técnica de la electroforesis<sup>4</sup>. Por regla general las dos técnicas electroforéticas más utilizadas son la electroforesis capilar y la electroforesis en gel de agarosa.

---

<sup>2</sup> SSR es el acrónimo en inglés de *Simple Sequence Repeats*. Las secuencias de tipo microsatélite (SSR) están constituidas por unidades cortas de 1 a 6 pares de bases, que se repiten en tándem un elevado número de veces. Cada secuencia SSR se define por el tipo de unidad repetida y por el sitio que ocupan en el genoma (locus) (González, citado por Picó & Esteras, 2011).

<sup>3</sup> La PCR es una técnica que consiste en obtener un gran número de copias de un fragmento de ADN, partiendo de una mínima cantidad del fragmento original o molde. Para replicar las hebras de ADN se utilizan ciclos de altas y bajas temperaturas, que sirven para separar las hebras, crear nuevas y volver a unir las (polimerización) para poder duplicarlas. Es una técnica que ofrece una replicación exponencial de los fragmentos de ADN que deseamos ampliar, obteniendo millones de copias del fragmento base.

<sup>4</sup> La electroforesis es una técnica para la separación de moléculas según la movilidad de éstas en un campo eléctrico.

ACCACACACACATT	(CA) <sub>5</sub>
ACCACACACA--TT	(CA) <sub>4</sub>
ACCACACA-----TT	(CA) <sub>3</sub>
ACCACA-----TT	(CA) <sub>2</sub>
ACCA-----TT	(CA) <sub>1</sub>

Figura 2 – Ejemplo de una secuencia de repetición breve. Al tipo de microsatélites de este ejemplo se les denomina perfectos y son repeticiones en un número mínimo de unidades sin interrupción y sin repeticiones adyacentes Fuente: <https://wiki.uiowa.edu/pages/viewpage.action?pageId=127692770>

La electroforesis capilar realiza la separación de los fragmentos de ADN en un capilar de sílice fundida de diámetro muy pequeño (10-200 µm) y los resultados se obtienen mediante el marcaje del DNA con múltiples fluorocromos que se añaden a los cebadores hacia delante (*primers forward*)<sup>5</sup>. Mediante esta técnica se pueden analizar hasta 4 marcadores a la vez, ya que el dispositivo es capaz de distinguir 4 colores (verde, azul, amarillo y rojo). Con lo que en una placa de 96 pocillos podemos realizar 384 análisis. Los resultados se visualizan en un electroferograma (Figura 3) que nos indica los alelos de cada muestra analizada según el marcador utilizado.

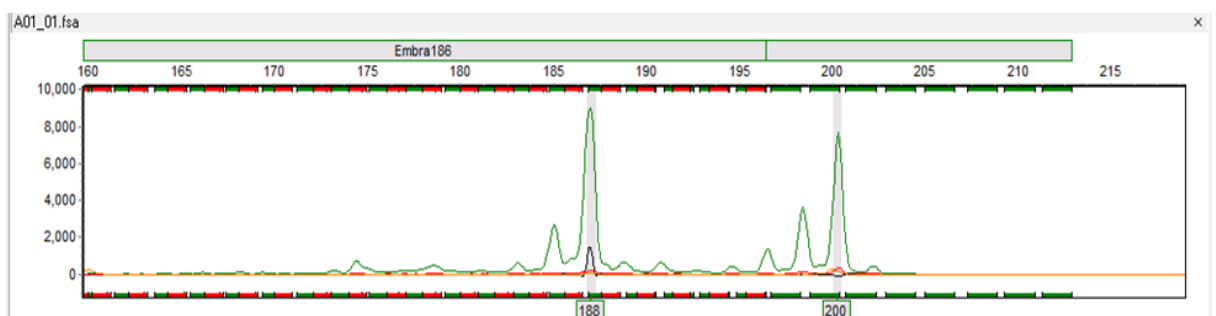


Figura 3 - Ejemplo de electroferograma analizado mediante el programa GeneMarker (Softgenetics). Fuente: propia.

La electroforesis en gel de agarosa separa los fragmentos de ADN por su tamaño y carga. Para visualizar los resultados se hace un gel, con un agente intercalante (se usa como aclarador de ácidos nucleicos), que dispone de unos pocillos para insertar una mezcla de ADN (agua HPLC<sup>6</sup>, ADN amplificado mediante PCR y azul de bromofenol). El gel se deposita en una cubeta de electroforesis con una solución tampón, se carga la solución de ADN y se aplica una corriente eléctrica. Los fragmentos migran del ánodo al cátodo, debido a la carga negativa del

<sup>5</sup> Los *primers* son secuencias de ADN complementarias a uno de los extremos de cada una de las dos hebras del ADN. Si la hebra va de 3' a 5' del ADN (secuencia AGCT) se le engancha un *primer forward* que irá de 5' a 3'. Si la hebra va de 5' a 3' (secuencia TCGA) se le engancha un *primer reverse* que irá de 3' a 5'.

<sup>6</sup> HPLC: Cromatografía Líquida de Alta Eficacia.

ADN. Una vez finalizado se deposita el gel en una cámara con transluminador de luz UV y se toma una fotografía (Figura 4).

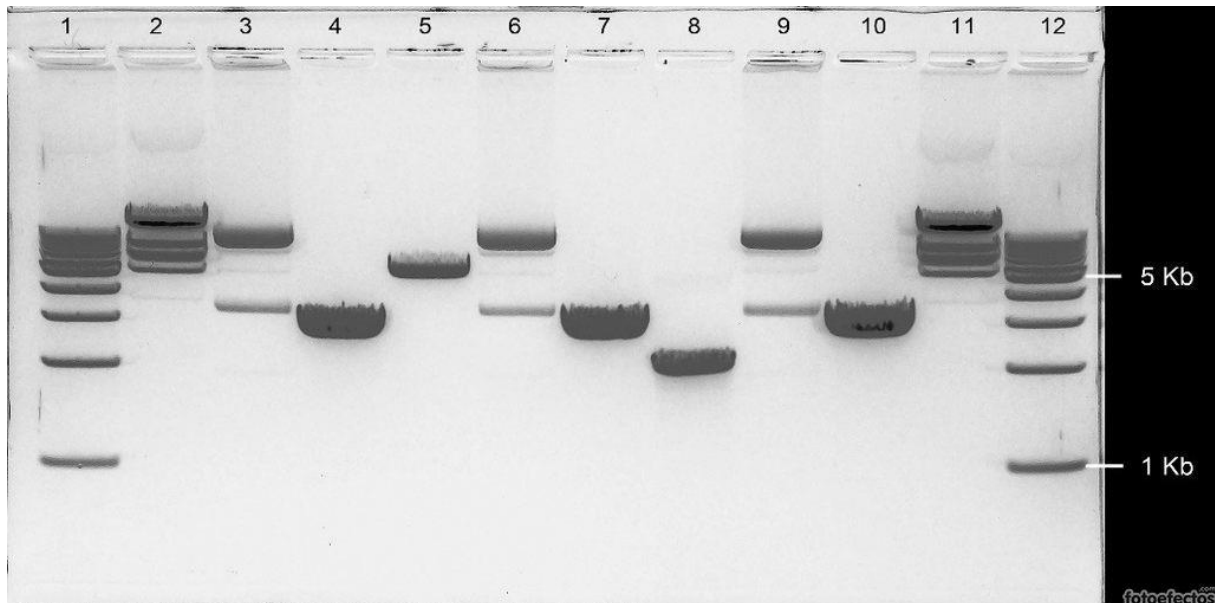


Figura 4 – Ejemplo de resultados obtenidos mediante gel de agarosa. Los fragmentos se separan en función de su tamaño y se ordenan de mayor (5 Kb) a menor tamaño (1 Kb). Kb hace referencia a Kilobase o lo que es lo mismo mil pares de bases.

Entre los marcadores disponibles hoy día, los más prometedores son los SNPs (*single nucleotide polymorphism*). Estos se pueden genotipar utilizando un gran abanico de técnicas, entre las que se incluyen las placas de matriz dinámicas de circuito fluídico integrado (dynamic array - *Integrated Fluidic Circuit*) permite el análisis simultáneo de los marcadores con las muestras de ADN. En el mismo soporte se realiza la PCR cuantitativa en tiempo real, denominada qRT-PCR, basada en sondas TaqMan<sup>7</sup>. Esta técnica representa además un ahorro en reactivos, ya que las mezclas se realizan a escala nanométrica, y también tendremos menos pasos de pipeteo (reduciendo los errores humanos). Para tener una idea de la magnitud de esta herramienta, en una placa de 96x96 se pueden hacer un total de 9.216 reacciones en un tiempo de unas 4 horas, tiempo que incluye el preparado de la placa más el análisis.

Para el género *Prunus* existen unos 103.342 marcadores (<https://www.rosaceae.org>).

---

<sup>7</sup> Las sondas TaqMan son sondas de hidrólisis diseñadas para incrementar la especificidad de la PCR cuantitativa. El método fue descrito por Holland, et al. (1991) y actualmente es una marca registrada de Roche Molecular Systems.

## 1.2.2 Mapas de ligamiento genético

Con los marcadores moleculares podemos construir mapas de ligamiento genético (Figura 5), que consiste en una representación del orden lineal de un grupo de marcadores a lo largo del genoma. La distancia que hay entre *loci* se da en unidades  $\text{cM}^8$  (centimorgans) y es proporcional a la frecuencia de recombinación entre los mismos. “Los marcadores ordenados se distribuyen en grupos de ligamiento, cuyo número debe corresponder al número haploide de cromosomas de la especie cuando el mapa tiene una cobertura genómica adecuada” (González, 2008).

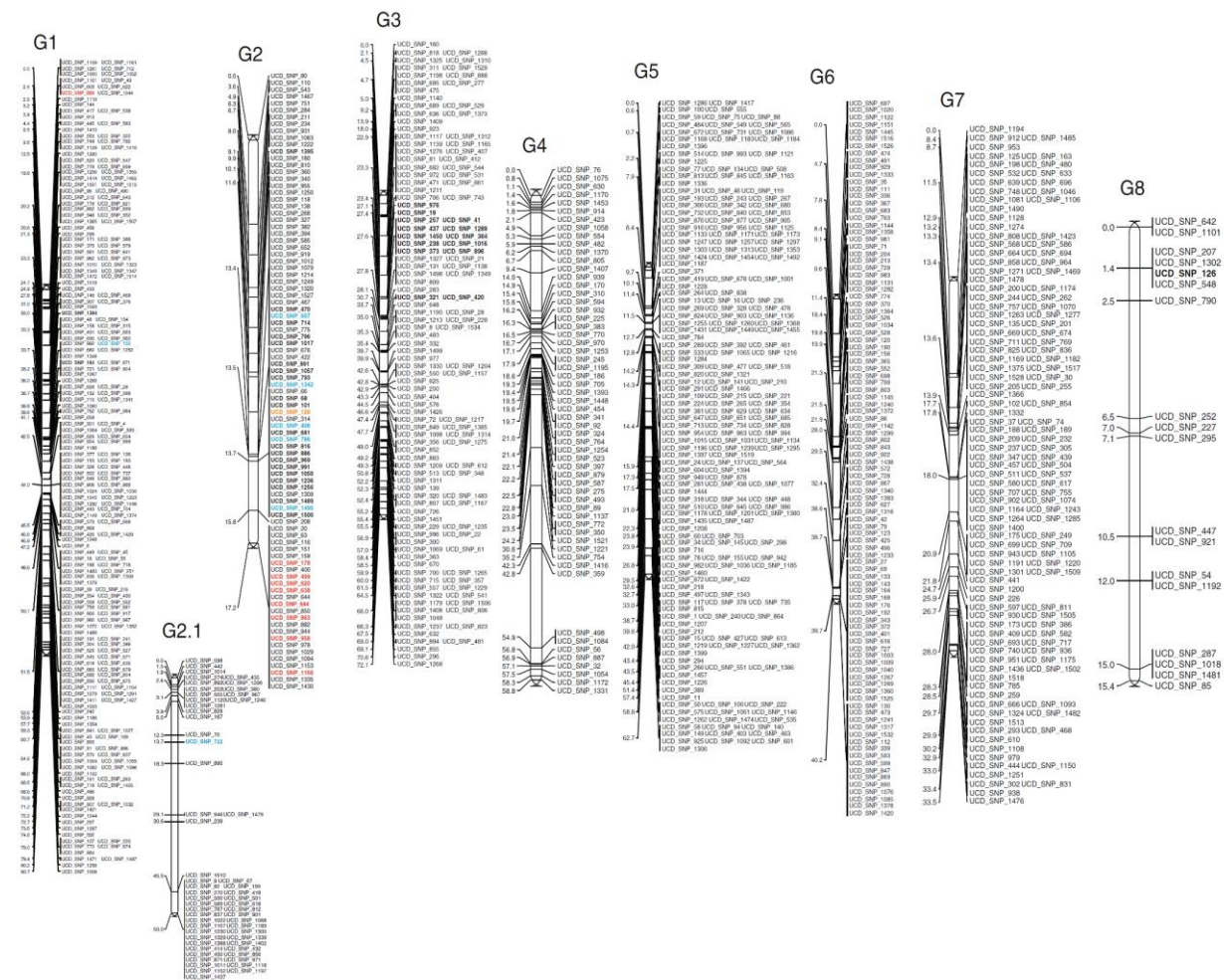


Figura 5 - Ejemplo de un mapa SNP de creado usando JoinMap®4.0. Los números en el lado izquierdo muestra la distancia en cM desde la parte superior de cada cromosoma. Fuente: (Martínez-García, et al., 2013)

<sup>8</sup> Un centimorgan (abreviado cM) es la unidad de los mapas genéticos de ligamiento realizados por recombinación en eucariotas diploides. Se trata de una unidad indivisible: el prefijo centi-, en este caso, no significa centésima parte de Morgan. Equivale a un 1% de frecuencia de recombinación, o lo que es lo mismo, a 1 unidad de mapa (Wikipedia, 2015).

### **1.2.3 Ejemplo de selección asistida por marcadores: maduración lenta o *slow ripening* (SR)**

El melocotón presenta en su genoma una mutación llamada maduración lenta o *slow ripening* (SR), que evita el proceso de maduración normal. Los individuos que presentan este fenotipo se descartan en los programas de mejora de melocotón.

El rasgo SR está determinado por un solo gen (Sr / sr), dónde el homocigoto recesivo (sr / sr) confiere el fenotipo SR (Meneses, et al., 2016).

Eduardo et al. (2015) identificaron una gran delección, en el cromosoma 4 (Pp04), alrededor del gen ppa008301 en individuos SR, lo que fue confirmado por Nuñez-Lillo et al. (2015). La eliminación abarcó una región de 26.6 KB entre las posiciones 11.111.981 y 11.137.943 del genoma del melocotón, versión 2.0, y contenía dos genes: ppa008301m (ANAC072) y ppa021959m (putativa transposasa). Basado en la secuencia de esta región, Eduardo et al. (2015) y Nuñez-Lillo et al. (2015) identificaron dos marcadores, uno codominante que co-segregó con el fenotipo SR en todos los materiales probados y otro dominante que sirve para identificar individuos SR pero no los heterocigotos, cosa que si permite el codominante. Este marcador fue validado en las variedades 'Belbinette' x 'Nectalady' y en las poblaciones F<sub>2</sub> de la variedad 'Venus', llegando a la conclusión de que esta herramienta es útil para detectar individuos SR en melocotonero.

El detectar esta mutación con un marcador codominante hace que se pueda descartar todos los individuos que llevan en su genoma la mutación SR y evitar los costes asociados a plantar individuos no productivos (ocupación de terreno, agua de regadío, fertilizantes, tratamientos fitosanitarios y costes asociados a la mano de obra).

### **1.2.4 Nuevas estrategias de mejora basadas en la selección asistida por marcadores: La Resíntesis**

La Resíntesis (Arús et al. 2016) es una nueva estrategia de mejoramiento basada en marcadores moleculares, que permite identificar en la progenie autofecundada de un cultivar:

- a) Un conjunto de líneas con la mayoría de su genoma como en el cultivar original y el resto homocigotos para uno de los dos haplotipos (fragmentos fijos en heterocigosis de fondo, FIHB). Necesita una generación F<sub>2</sub>. Esta estrategia se denomina Resíntesis A (Figura 6).
- b) Dos líneas complementarias que, cuando se cruzan, generarán nuevamente una leve versión diferente de la variedad original. Necesita tres generaciones. Esta técnica se denomina Resíntesis B (Figura 7).

Como base genética se ha de tener en cuenta dos aspectos:

- 1) El número de eventos de recombinación por meiosis por cromosoma es pequeño (1-2, rara vez más).
- 2) Es posible extraer cromosomas intactos no recombinados a una tasa relativamente alta: 1/4 suponiendo una recombinación por cromosoma, 1/8 si se producen dos cruces por cromosoma (probabilidad de 1/8 o 1/32 por cromosoma para obtener pares homocigotos no recombinados).

Según los cálculos estadísticos de Arús et al (2016) se deberían obtener alrededor de un 3% de individuos con una cantidad  $\geq 5$  de cromosomas intactos heterocigotos (IHC).

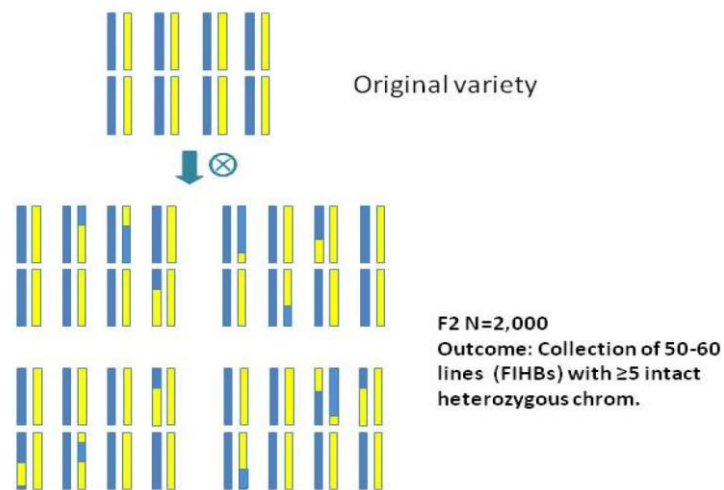


Figura 6 - Generación de una colección de individuos casi idéntica a una variedad de melocotón, pero con pocos segmentos en homocigosis (el conjunto de FIHB). Todos los individuos del conjunto presentado tienen cinco cromosomas heterocigotos como en el cultivar original y solo tres fragmentos homocigotos cada uno.

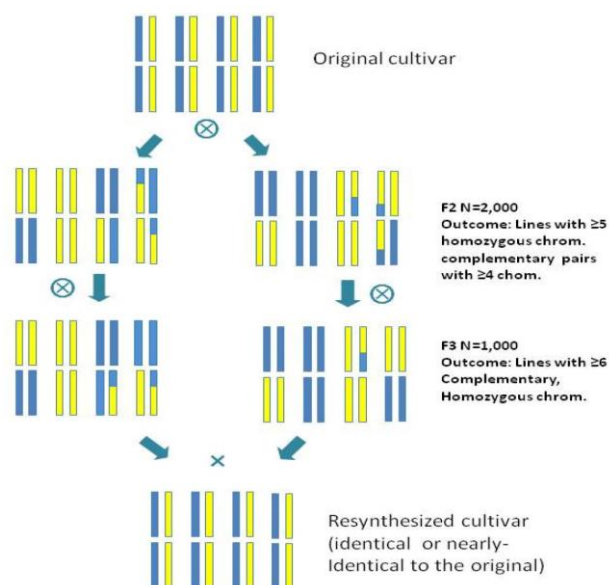


Figura 7 - Esquema de la resíntesis de un cultivar de melocotón. Los cromosomas homólogos del mismo par están en diferentes colores.

## **2 OBJETIVOS**

Este trabajo de fin de máster está incluido en el desarrollo de una prueba de concepto sobre la nueva estrategia de la Resíntesis, desarrollada por Arús, et al (2016). De forma específica en este máster se ha llevado a cabo el genotipado y la selección asistida por marcadores.

### **3 MATERIALES Y MÉTODOS**

#### **3.1 Material vegetal**

Durante 3 años (2015, 2016, 2017) se recolectaron los frutos obtenidos, mediante polinización abierta, de diferentes árboles de la variedad Sweet Dream (SD), ubicados en un huerto comercial, en la provincia de Lleida.

En 2015, después de la recolección de la fruta, se extrajeron aproximadamente 2000 embriones de los huesos de los melocotones y se siguió el protocolo de estratificación, que consiste en guardar las semillas durante 12 semanas a una temperatura de 4°C. Transcurrido ese tiempo, quedaron en buenas condiciones unas 600 semillas, que estaban en condiciones de plantarse en sustrato para posteriormente ser trasladadas a la estación experimental del IRTA de Gimènells, en la provincia de Lleida.

En 2016, se cosecharon 1600 melocotones. 800 embriones siguieron el mismo proceso de germinación del año anterior y se consiguió sembrar 737 semillas, obteniéndose un total de 136 plántulas. Con las 800 semillas restantes se utilizó un protocolo de rescate de embriones. De los 579 embriones que se pudo extraer se obtuvieron un total de 73 plántulas.

En 2017, se decidió dejar las frutas en exceso en los árboles para que los embriones tuvieran el tiempo suficiente para desarrollarse. 407 frutos fueron recogidos maduros en el árbol, 294 ya estaban en el suelo, pero no estaban podridos y 592 ya estaban podridos en el suelo. Finalmente se obtuvieron 77 plántulas.

Para el estudio de este trabajo se dispuso de un total de 418 plántulas de melocotonero (*Prunus persica*) de la polinización abierta de la variedad Sweet Dream (SD).

#### **3.2 Extracción de ADN**

El ADN nuclear se extrajo de las hojas más jóvenes mediante el protocolo de extracción Doyle-Doyle (1990) sin realizar el tratamiento final de la RNasa. El ADN extraído de todas las muestras disponibles se dispuso en placas PCR de 96 pocillos, de perfil estándar y transparente; preservándolas a -20°C, en un congelador.



### 3.3 Genotipado mediante SNPs (Single Nucleotide Polymorphism)

Para obtener el genotipado se utilizaron placas de matriz dinámica (*Dynamic Array*) de 96x96 de Fluidigm, de circuito fluidoico integrado. Las placas permiten cargar los 96 marcadores, con sus cebadores o *primers*, y se hacen reaccionar con 96 muestras, obteniendo un total de 9.216 reacciones. En cada placa se incluían el parental SD y un control negativo sin muestra de ADN.

Antes de analizar las placas se prepararon las muestras con el protocolo KASPar Assay Primer Mix (Fluidigm, 2018). En este proceso se añaden los cebadores específicos de alelo directos (*allele specific primer*), el cebador reverso (*common reverse primer*) y la solución tampón (*DNA suspensión buffer*).

El análisis de las placas se realizó mediante la plataforma Fluidigm BioMark HD System, modelo IFC Controller HX. Los resultados se analizaron mediante el software Fluidigm SNP Genotyping Analysis, versión 4.1.3.

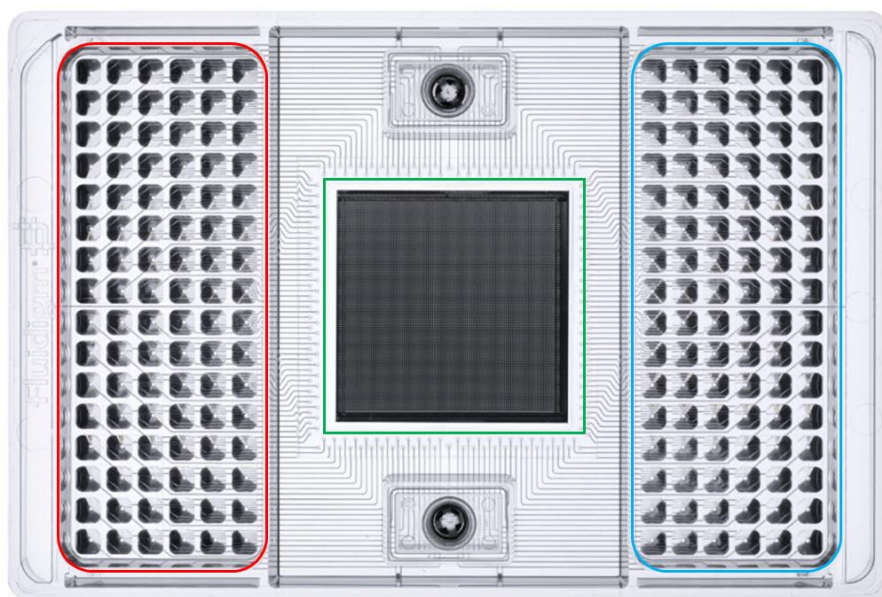


Figura 8 - Placa de matriz dinámica (*Dynamic Array*) de 96x96 de Fluidigm. En la zona marcada en rojo hay 96 pocillos dónde se cargan las muestras, en la zona marcada en azul hay 96 pocillos dónde se cargan los marcadores y en la zona marcada en verde es dónde se dan las reacciones de los 96 individuos con los 96 marcadores.



Figura 9 - Plataforma Fluidigm BioMark HD System, modelo IFC Controller HX.

Fuente: <https://www.labx.com/item/fluidigm-biomark-hd-real-time-pcr-w-ifc-hx-mx-controller/4250247>

### 3.4 Genotipado mediante microsatélites (*SSR: Simple Sequence Repeats*)

Para la amplificación del ADN, mediante la reacción en cadena de la polimerasa (PCR), se utilizó un termociclador ProFlex™ 96-well PCR System de la empresa Applied Biosystems. En cada placa de 96 también se incluían el parental SD y un control negativo sin muestra de ADN. Los marcadores SSR, fueron analizados con un secuenciador capilar Applied Biosystems 3130 Genetic Analyzer. Finalmente, se obtuvieron los resultados con el software **GeneMapper 5** de Applied Biosystems y **GeneMarker v1.85** de SoftGenetics.

Los reactivos utilizados para la PCR fueron (mezcla 1x)<sup>9</sup>: 5.9µl de agua HPLC<sup>10</sup>, 1µl de Lab1X tampón (Lab 10X: KCl 50mM, Tris-HCl 10mM pH 8.3, gelatina 0.0001%), 0.3µl MgCl<sub>2</sub> 50mM, 0.2µl dNTP 10mM, 0.2µl de cebador directo marcado con fluorocromo 0.2µM, 0.2µl cebador reverso 0.2µM, 0.2µl de Taq polimerasa 5U/µl y 2 µl de ADN 20ng/µl. El programa de amplificación se estableció de la siguiente manera: 1 min a 94°C; 35 ciclos de 15 s a 94°C, 20 s a *Ta*<sup>11</sup>°C y 30 s a 72°C; finalmente 5 min a 72°C.

<sup>9</sup> La mezcla es la cantidad necesaria para la PCR de 1 individuo, para el análisis de x muestras se tiene que multiplicar el volumen por la cantidad total de x muestras.

<sup>10</sup> HPLC: Cromatografía líquida de alta presión.

<sup>11</sup> La *Ta* es la temperatura de *annealing* (recocido) de cada primer. La *Ta* del marcador CPPCT029 es de 55°C y la del marcador CPPCT21395 es de 57°C.

Después de la amplificación se prepararon las muestras, para ser analizadas mediante electroforesis capilar, con los siguientes reactivos (mezcla 1x): 12µl de formamida Hi-Di (*highly deionized*), 0.35µl de marcador de peso molecular **LIZ** (*Internal Lane Size Standard*), 2µl de la PCR CPPCT029 y 2µl de la PCR CPPCT21395, se desnaturalizan las muestras a 94°C durante 3 minutos. Por último, se cargaron en el secuenciador ABI

### 3.5 Genotipado del marcador SRcod

Para la amplificación del ADN, mediante la reacción en cadena de la polimerasa (PCR), se utilizó un termociclador ProFlex™ 96-well PCR System de la empresa Applied Biosystems. Las muestras se analizaron utilizando la electroforesis en gel de agarosa. Finalmente, se obtuvo la imagen del gel, mediante visualización con luz UV, para ver que individuos presentan la banda *sr/sr*.

Los reactivos utilizados para la PCR fueron (mezcla 1x): 5.5µl de agua HPLC, 1µl de Lab1X tampón (Lab 10X: KCl 50mM, Tris-HCl 10mM pH 8.3, gelatina 0.0001%), 0.3µl MgCl<sub>2</sub> 50mM, 0.2µl dNTP 10mM, 0.2µl del cebador directo (NR1-F) 10µM, 0.2µl del cebador reverso (NR1-R) 10µM, 0.2µl del cebador directo (MDPr2-F) 10µM, 0.2µl del cebador reverso (MDPr2-R) 10µM, 0.2µl de Taq polimerasa 5U/µl y 2 µl de ADN 20ng/µl. El programa de amplificación se estableció de la siguiente manera: 1 min a 94°C; 35 ciclos de 15 s a 94°C, 30 s a 51°C y 30 s a 72°C; finalmente 5 min a 72°C.

Los productos de la PCR se separaron en geles de agarosa al 1.5% en tampón TAE<sup>12</sup> 1x y se tiñeron con bromuro de etidio (20 µl BrEt/100 ml de solución tampón). Los pocillos del gel se cargaron con una solución de: 8 µl de H<sub>2</sub>O HPLC, 2 µl de ADN 20ng/µl y 2 µl de azul de bromofenol (0.02%).

### 3.6 Tratamiento previo de los datos antes de realizar el mapa genético y la Resíntesis

Una vez se dispone de todo el genotipado analizado, mediante SNPs, microsatélites y el marcador SRcod, se deben descartar aquellos marcadores que no segregaron, no amplificaron o no se distinguieron las clases genotípicas.

---

<sup>12</sup> TAE: Solución tampón compuesta por: **Tris** (Tris-(hidroximetil)-aminometano), **Ácido Bórico/Ácido Acético**, **EDTA** (Ácido Etilendiaminotetraacético).

Los individuos analizados mediante SNPs y microsatélites presentan datos perdidos (*missing values*) en algunas muestras. Según Dagnino (2014) el sesgo que introduce o puede introducir la falta de datos es proporcional al número de pérdidas, con lo que trabajar con un 10% de valores perdidos es inaceptable.

Existen tres estrategias para lidiar con la problemática de los datos faltantes:

1. Omitir variables con datos faltantes.
2. Omitir individuos en quienes hay datos faltantes.
3. Estimar (imputar) los datos faltantes donde estos son reemplazados con valores predichos desde los datos presentes.

En este trabajo nos regimos por el criterio de no trabajar con individuos que presenten más de un 10% de datos perdidos.

Los datos seleccionados se disponen en una hoja de cálculo introduciendo todos los marcadores aptos en filas, ordenados por número y posición de cada cromosoma, y todos los individuos válidos ordenados por columnas; obteniendo una matriz de **m** (marcadores ya seleccionados) x **n** (individuos con un porcentaje de datos perdidos <10%), en adelante se denominará a esta matriz como **M1**.

### 3.7 Construcción de un mapa de ligamiento genético

Con los datos de la matriz M1 se elaboró el mapa de ligamiento genético mediante el software **JoinMap** v4.1 de la empresa **Kyazma**, para una población de tipo F<sub>2</sub>. El programa nos da la cantidad de grupos posibles en función de la puntuación LOD<sup>13</sup>, que hace referencia al logaritmo de las probabilidades de que dos genes o loci se encuentren ligados y por lo tanto se heredan unidos con más frecuencia de lo habitual. En nuestro caso se escogieron los grupos con una puntuación de LOD igual a 4 para obtener los 8 grupos posibles, tantos como cromosomas tiene la especie de estudio. Para la función de mapa se usó la función de Kosambi<sup>14</sup> (1944).

---

<sup>13</sup> **LOD**: Acrónimo en inglés de *logarithm of the odds*. LOD = log(probabilidad de ligamiento con frecuencia de recombinación R/probabilidad de transmisión no ligada).

<sup>14</sup> La **función de Kosambi** asume una disminución lineal de la interferencia a medida que aumenta la distancia genética ( $i = 1 - 2r$ ); siendo  $r$  la fracción de recombinación y  $0 \leq r \leq 0.5$ .

## 3.8 Resíntesis

### 3.8.1 Metodología de estudio de la Resíntesis A

El objetivo de la Resíntesis A es encontrar, a partir de una población F<sub>2</sub>, individuos con una configuración genética similar a la variedad original autofecundada, que en nuestro caso es la variedad Sweet Dream. Consideramos similares individuos con pocos segmentos en homocigosis y a poder ser, cinco o más cromosomas totalmente heterocigotos, idénticos a los del cultivar original. Para lograr encontrar los individuos que se ajustan a esta condición se debe proceder al cálculo de la heterocigosidad y escoger aquellos individuos que presenten los valores más elevados.

Con la matriz M1 se realizará el cálculo de la heterocigosidad ordenando los individuos de mayor a menor heterocigosidad, eligiendo sólo aquellos individuos con mayor heterocigosidad.

#### 3.8.1.1 Cálculo de la heterocigosidad

Con los datos de la matriz M1 se obtiene la suma de los alelos heterocigotos observados (XY), la suma de los alelos homocigotos observados (XX o YY) y la suma de los datos perdidos, para cada individuo, y se calcula la heterocigosidad mediante la siguiente ecuación:

$$\text{Heterocigosidad} = \frac{\sum_{i=1}^m (\text{datos de heterocigotos observados para cada locus})}{(m - \text{missing values})} = \frac{\sum_{i=1}^m (XY)}{(m - \text{missing values})}$$

Dónde:

m = cantidad total de marcadores válidos utilizados

La suma total de alelos homocigotos, heterocigotos y de los datos perdidos debe ser igual a la cantidad de marcadores válidos utilizados.

### 3.8.2 Metodología de estudio de la Resíntesis B

La finalidad de la Resíntesis B es, a partir de una población F<sub>2</sub>, encontrar parejas de individuos complementarios que al cruzarse produzcan individuos muy similares a la línea original autofecundada. Se consideran complementarios individuos que si uno cuenta con el alelo homocigoto XX el otro ha de contar con el alelo homocigoto YY, de tal manera que cuando se crucen obtengamos un 100% de individuos heterocigotos XY.

Con la matriz M1 se realizará el cálculo de la heterocigosidad ordenando los individuos de menor a mayor heterocigosidad, eligiendo sólo aquellos individuos con menor heterocigosidad.

Suponiendo que tenemos un individuo A que se cruza con un individuo B, las regiones pueden presentar 4 tipos de emparejamientos:

**Emparejamiento complementario**, el individuo A presenta los alelos XX y el individuo B los alelos YY. Al cruzarse obtendremos el 100% de heterocigotos XY. Este es el tipo de emparejamientos que pretendemos buscar, ya que a más cantidad de regiones complementarias más parecido tendremos con la variedad parental.

**Emparejamiento de 1 alelo homocigoto con 1 heterocigoto**, el individuo A presenta los alelos XX y el individuo B presenta los alelos XY. Al cruzarse obtendremos un 50% de homocigotos XX y un 50% de heterocigotos XY.

**Emparejamiento de 2 alelos heterocigotos**, el individuo A presenta los alelos XY y el individuo B presenta los alelos XY. Al cruzarse obtendremos un 25% de homocigotos XX, un 50% de heterocigotos XY y un 25% de homocigotos YY.

**Emparejamiento de 2 alelos homocigotos**, el individuo A presenta los alelos XX y el individuo B presenta los alelos XX. Al cruzarse obtendremos un 100% de homocigotos XX. Seleccionaremos individuos con el menor número posible de este tipo de emparejamientos.

### 3.8.3 Método para encontrar los individuos aptos para realizar cruzamientos

Como aún no se dispone de un método bioinformático para encontrar que parejas son las más adecuadas, se ha de hacer una correlación estadística de todos los individuos, con un valor de heterocigosidad  $\leq 0,30$ , mediante el coeficiente de correlación de Pearson. Al tratarse de datos no numéricos se asigna un valor numérico en función de las X que presentaba cada locus, así pues, si presenta una XX se le asigna un valor de 2, si presenta una XY se le asigna un valor de 1 y si presenta una YY se le asigna un valor de 0, tal como se puede ver en el ejemplo de la Tabla 1. El coeficiente de correlación puede dar un rango de valores de -1 a 1. Como se asigna un valor de 2 a XX y de 0 a YY debemos esperar que los individuos que presenten mejor complementariedad serán aquellos que presenten un coeficiente de correlación negativo, ya que en la medida que aumenta una variable disminuye la otra. A mayor correlación negativa más regiones complementarias se presentarán y se obtendrán las parejas de candidatos más adecuados para realizar un cruzamiento que de un individuo parecido al parental original Sweet Dream.

Tabla 1 – Ejemplo de conversión de variables no numéricas a numéricas. Se asigna un valor numérico en función de las X que presenta cada locus, quedando: XX=2, XY=1, YY=0.

Datos originales				Datos transformado numéricos			
Individuo a	Individuo b	Individuo c	Individuo d	Individuo a	Individuo b	Individuo c	Individuo d
XX	XX	XX	XY	2	2	2	1
XX	XY	XX	XY	2	1	2	1
XX	XY	XX	YY	2	1	2	0
XX	XY	XX	YY	2	1	2	0
XY	YY	XY	YY	1	0	1	0
XX	YY	XY	XY	2	0	1	1
YY	XY	YY	XX	0	1	0	2
YY	XY	YY	XX	0	1	0	2

## **4 RESULTADOS**

### **4.1 Material vegetal disponible para los ensayos**

Para el análisis de la Resíntesis se obtuvieron y germinaron 2.617 semillas procedentes de la autofecundación de la variedad SD. De los 2.617 individuos plantados se obtuvieron un total de 418 plántulas viables, lo que nos indica que la tasa de germinación de este experimento está en torno al 16 %.

### **4.2 Genotipado**

#### **4.2.1 Genotipado mediante SNP**

Los SNPs analizados se seleccionaron, antes del inicio de este master, mediante el análisis de la resecuencia de la variedad Sweet Dream. Utilizando las resecuencias primero se hizo una anotación de los SNPs heterocigotos y posteriormente se observó su distribución en el genoma para saber que regiones estaban fijadas, es decir, eran homocigotas. Los criterios para la selección fueron que estuvieran en regiones heterocigotas y que cubrieran el máximo de genoma posible. Finalmente se seleccionaron los 96 SNPs que se presentan en la Tabla 2.

La posición física de los marcadores extrajo de la versión 2.0 del genoma del melocotón (Verde, et al., 2017).

Se ensayaron un total de 5 placas de 96 x 96 para poder analizar las 418 muestras. Mediante el software de Fluidigm se pudieron exportar los resultados a una hoja de cálculo, en forma matricial de 96 x 96 (marcadores x individuos) tal como los que se muestran en la Tabla 3.



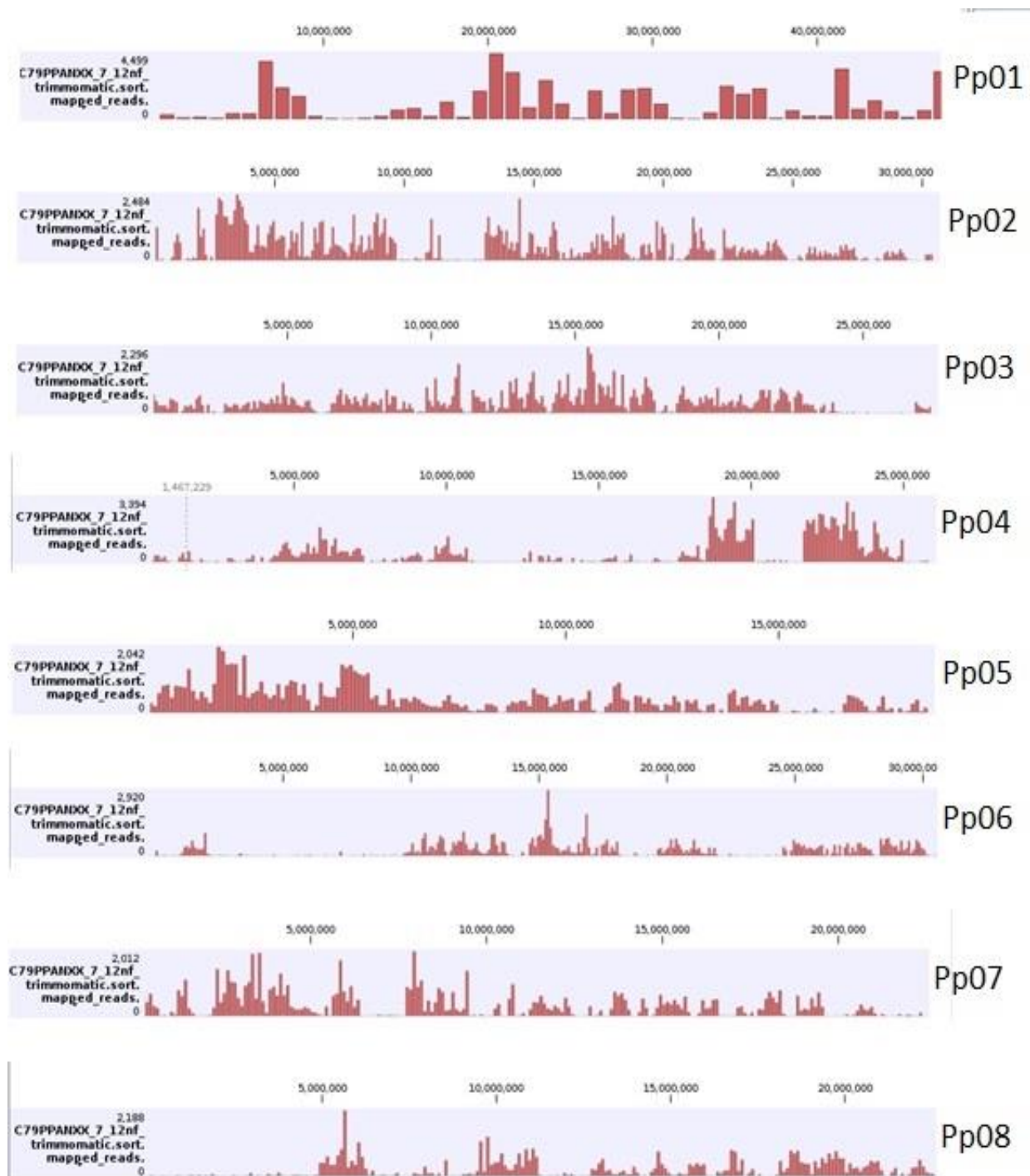


Figura 10 - Mapa de la resecuenciación del cultivar SD, dónde se muestran todas las regiones de homocigosidad (zonas sin barras) y heterocigosidad (zonas con alta concentración de barras). En el eje de las X se representan los pares de bases y en el eje de las Y el número de SNPs heterocigotos. Las zonas de mayor heterocigosidad es dónde interesa escoger los marcadores, ya que son las zonas dónde tendremos segregación.

**Tabla 2 - Marcadores utilizados para el análisis de las muestras mediante SNP.**

<b>Orden genoma melocotón V 2.0</b>	<b>Marcador</b>	<b>Cromosoma</b>	<b>Región</b>	<b>Orden genoma melocotón V 2.0</b>	<b>Marcador</b>	<b>Cromosoma</b>	<b>Región</b>
1	SNP_Pp01-8080	Pp01	8.080	49	SNP_IGA_528949	Pp04	24.187.959
2	SNP_IGA_2651	Pp01	883.576	50	SNP_IGA_544495	Pp05	610.569
3	SNP_IGA_17419	Pp01	6.000.280	51	SNP_IGA_548037	Pp05	1.376.475
4	SNP_IGA_19514	Pp01	6.872.054	52	SNP_IGA_553456	Pp05	2.477.309
5	SNP_IGA_31646	Pp01	12.187.249	53	SNP_IGA_571548	Pp05	5.485.331
6	SNP_Pp01-15435423	Pp01	15.435.423	54	SNP_IGA_572589	Pp05	5.813.029
7	SNP_IGA_67137	Pp01	20.219.128	55	SNP_IGA_585810	Pp05	9.347.092
8	SNP_IGA_79455	Pp01	23.529.157	56	SNP_IGA_591439	Pp05	11.196.070
9	SNP_IGA_86968	Pp01	26.640.558	57	SNP_IGA_595126	Pp05	12.682.824
10	SNP_IGA_103771	Pp01	33.744.976	58	SNP_IGA_596393	Pp05	13.411.012
11	SNP_IGA_120784	Pp01	41.413.142	59	SNP_IGA_600493	Pp05	14.879.113
12	SNP_IGA_125219	Pp01	46.984.218	60	SNP_IGA_602605	Pp05	16.636.368
13	SNP_IGA_123719	Pp01	47.532.400	61	SNP_IGA_616458	Pp06	1.304.512
14	SNP_IGA_238135	Pp02	241.474	62	SNP_Pp06-2176846	Pp06	2.176.846
15	SNP_IGA_136625	Pp02	2.002.076	63	SNP_IGA_610889	Pp06	3.034.364
16	SNP_IGA_148760	Pp02	3.109.681	64	SNP_Pp06-7216093	Pp06	7.216.093
17	SNP_IGA_195134	Pp02	6.545.323	65	SNP_IGA_627328	Pp06	7.550.350
18	SNP_IGA_218801	Pp02	9.053.696	66	SNP_IGA_639062	Pp06	11.085.127
19	SNP_IGA_242118	Pp02	13.717.350	67	SNP_IGA_640395	Pp06	11.298.491
20	SNP_IGA_252957	Pp02	15.849.694	68	SNP_IGA_655825	Pp06	14.681.941
21	SNP_IGA_258854	Pp02	17.219.077	69	SNP_IGA_663057	Pp06	17.959.563
22	SNP_IGA_266063	Pp02	19.744.284	70	SNP_IGA_671806	Pp06	21.011.215
23	SNP_IGA_275502	Pp02	22.301.505	71	SNP_IGA_682254	Pp06	24.511.760
24	SNP_IGA_283325	Pp02	25.849.495	72	SNP_IGA_691624	Pp06	27.207.318
25	SNP_IGA_285058	Pp02	26.895.362	73	SNP_IGA_696341	Pp06	28.500.652
26	SNP_IGA_898654	Pp03	742.389	74	SNP_IGA_700653	Pp06	30.012.439
27	SNP_Pp03-3917979	Pp03	3.917.979	75	SNP_IGA_726222	Pp07	350.629
28	SNP_IGA_299796	Pp03	4.209.014	76	SNP_IGA_717591	Pp07	2.354.039
29	SNP_IGA_322698	Pp03	9.911.321	77	SNP_IGA_704075	Pp07	4.239.920
30	SNP_IGA_428056	Pp03	11.876.389	78	SNP_Pp07-5647370	Pp07	5.827.566
31	SNP_IGA_894039	Pp03	13.664.157	79	SNP_IGA_752104	Pp07	8.182.487
32	SNP_IGA_325850	Pp03	15.520.683	80	SNP_IGA_758767	Pp07	10.388.929
33	SNP_IGA_344612	Pp03	19.098.840	81	SNP_IGA_774557	Pp07	13.833.942
34	SNP_IGA_349233	Pp03	20.300.608	82	SNP_IGA_776826	Pp07	14.891.881
35	SNP_IGA_356179	Pp03	22.294.942	83	SNP_IGA_781003	Pp07	16.422.167
36	SNP_IGA_368926	Pp04	407.324	84	SNP_IGA_784777	Pp07	18.244.540
37	SNP_IGA_379856	Pp04	1.477.791	85	SNP_IGA_794167	Pp08	373.237
38	SNP_IGA_389204	Pp04	4.843.247	86	SNP_IGA_797492	Pp08	1.248.336
39	SNP_IGA_395253	Pp04	6.202.308	87	SNP_IGA_803699	Pp08	2.564.296
40	SNP_IGA_398213	Pp04	6.895.567	88	SNP_Pp08-3702593	Pp08	3.702.593
41	SNP_IGA_403741	Pp04	9.106.604	89	SNP_IGA_821894	Pp08	5.071.328
42	SNP_IGA_409453	Pp04	10.396.616	90	SNP_IGA_851849	Pp08	10.999.017
43	SNP_Pp04-13754607	Pp04	13.754.607	91	SNP_IGA_860815	Pp08	13.540.876
44	SNP_IGA_420819	Pp04	13.949.755	92	SNP_Pp08-13670362	Pp08	13.670.362
45	SNP_IGA_440116	Pp04	16.084.694	93	SNP_IGA_871082	Pp08	16.740.766
46	SNP_IGA_454568	Pp04	18.760.831	94	SNP_IGA_878044	Pp08	18.644.875
47	SNP_IGA_480804	Pp04	20.068.418	95	SNP_IGA_883524	Pp08	20.833.164
48	SNP_IGA_513502	Pp04	22.616.919	96	SNP_IGA_884755	Pp08	21.455.147

**Tabla 3 –Ejemplo de una muestra parcial, de la matriz de datos que entrega el software de Fluidigm. Los individuos se ordenan por columnas y los marcadores por filas. Existen 4 resultados posibles: XX (rojo) o YY (verde) si se trata de una región homocigota, XY (amarillo) si se trata de una región heterocigota y No Call (gris) que indica que se ha dado un problema de amplificación, segregación u otro tipo de problema y lo considera como dato no válido.**

<b>Marcador/Individuo</b>	<b>SDF2-01</b>	<b>SDF2-02</b>	<b>SDF2-03</b>	<b>SDF2-04</b>	<b>SDF2-05</b>	<b>SDF2-06</b>	<b>SDF2-07</b>	<b>SDF2-08</b>
SNP_Pp01-8080	XY	XY	XY	XY	XY	XY	XY	XY
SNP_IGA_2651	No Call	No Call	No Call	No Call	No Call	No Call	No Call	No Call
SNP_IGA_17419	XY	YY	YY	XY	YY	XX	XX	XY
SNP_IGA_19514	XY	XX	XX	XY	XX	YY	YY	XY
SNP_IGA_31646	YY	YY	YY	YY	YY	YY	YY	YY
SNP_Pp01-15435423	YY	YY	YY	YY	YY	YY	YY	YY
SNP_IGA_67137	XY	XX	XY	XX	XX	YY	YY	XY
SNP_IGA_79455	XY	YY	XY	YY	YY	XX	XX	XY
SNP_IGA_86968	XY	XX	XY	XX	XX	YY	YY	YY
SNP_IGA_103771	XX	XX	XX	XX	XX	XX	XX	XX
SNP_IGA_120784	YY	No Call	YY	No Call	YY	YY	XY	YY
SNP_IGA_125219	XX	YY	No Call	No Call	No Call	No Call	No Call	XX
SNP_IGA_123719	No Call	No Call	No Call	No Call	No Call	No Call	No Call	No Call
SNP_IGA_238135	No Call	No Call	No Call	No Call	No Call	No Call	No Call	No Call
SNP_IGA_136625	XY	XY	XY	XY	YY	XY	XY	XY
SNP_IGA_148760	XY	XY	XY	XY	XX	XY	XY	XY
SNP_IGA_195134	XY	XY	XY	XY	XX	XY	XY	XY
SNP_IGA_218801	XY	XY	XY	XY	YY	XY	XY	XY
SNP_IGA_242118	XY	XY	XX	XY	XX	XY	XY	XY
SNP_IGA_252957	XY	XY	XX	XY	XX	XY	XY	XY
SNP_IGA_258854	XY	XY	YY	XY	YY	XY	XY	YY
SNP_IGA_266063	XY	XY	XY	XY	YY	XX	XY	YY
SNP_IGA_275502	XY	XY	XY	XY	YY	XX	XY	YY
SNP_IGA_283325	YY	XX	XY	XY	XX	YY	XY	XX
SNP_IGA_285058	YY	XX	XY	XY	XX	XY	XY	XY

De los 96 marcadores SNP analizados se detectaron 34 que presentaron problemas de segregación, amplificación o distinción de las clases genotípicas (Tabla 4); Figura 11), con lo que se descartaron para la elaboración del mapa de ligamiento genético y el estudio de la Resíntesis.

Tabla 4-Marcadores descartados, ordenados por número de cromosoma y región. En la última columna se especifica el motivo por el cual fue descartado.

Marcador	Cromosoma	Región	Problema
SNP_Pp01-8080	Pp01	8080	No segrega/No amplifica
SNP_IGA_2651	Pp01	883576	No amplifica
SNP_IGA_31646	Pp01	12187249	No segrega
SNP_Pp01-15435423	Pp01	15435423	No segrega
SNP_IGA_103771	Pp01	33744976	No segrega
SNP_IGA_120784	Pp01	41413142	No segrega
SNP_IGA_125219	Pp01	46984218	No segrega/No amplifica
SNP_IGA_123719	Pp01	47532400	No amplifica
SNP_IGA_238135	Pp02	41160501	No segrega/No amplifica
SNP_Pp03-3917979	Pp03	742389	No segrega/No amplifica
SNP_IGA_299796	Pp03	3917979	No segrega
SNP_IGA_322698	Pp03	4209014	No segrega/No amplifica
SNP_IGA_428056	Pp03	9911321	No segrega/No amplifica
SNP_IGA_894039	Pp03	11876389	No segrega/No amplifica
SNP_Pp04-13754607	Pp04	10396616	No segrega
SNP_IGA_420819	Pp04	13754607	No segrega
SNP_IGA_616458	Pp06	16636368	No amplifica
SNP_Pp06-2176846	Pp06	1130915	No segrega
SNP_IGA_610889	Pp06	1304512	No amplifica
SNP_Pp06-7216093	Pp06	2176846	No segrega
SNP_IGA_627328	Pp06	3034364	No segrega
SNP_IGA_639062	Pp06	7216093	No segrega
SNP_IGA_671806	Pp06	14681941	No segrega
SNP_IGA_682254	Pp06	17959563	No segrega
SNP_IGA_726222	Pp07	28500652	No amplifica
SNP_IGA_717591	Pp07	30012439	No segrega/No amplifica
SNP_IGA_704075	Pp07	350629	No segrega/No amplifica
SNP_IGA_794167	Pp08	16422167	No se distinguen las clases genotípicas
SNP_IGA_797492	Pp08	18244540	No segrega
SNP_IGA_803699	Pp08	373237	No segrega/No amplifica
SNP_Pp08-3702593	Pp08	1248336	No segrega
SNP_IGA_860815	Pp08	5071328	No segrega
SNP_Pp08-13670362	Pp08	10999017	No segrega/No amplifica
SNP_IGA_883524	Pp08	16740766	No segrega

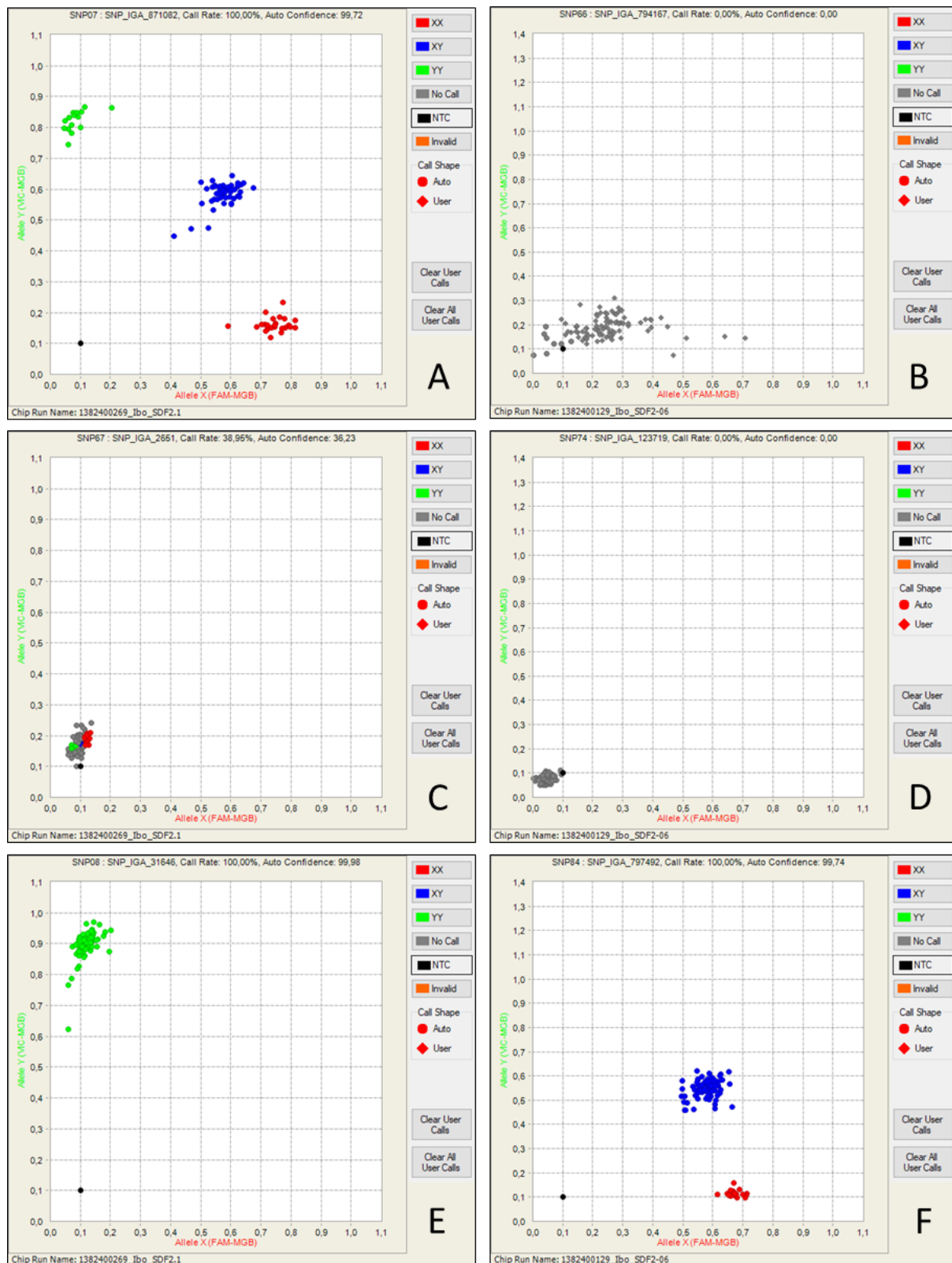


Figura 11 - La figura A corresponde al gráfico de dispersión del marcador SNP\_IGA\_871082 y se presenta como ejemplo de lo que es un marcador que segrega bien y que se considera apto (Co-dominante). La figura B corresponde al marcador SNP\_IGA\_794167 y se presenta como ejemplo de la imposibilidad de distinguir las clases genotípicas, aunque si tengamos segregación. Las figuras C y D corresponden a los marcadores SNP\_IGA\_2651 y SNP\_IGA\_123719 y se presenta como ejemplo de falta de amplificación, en este caso es imposible distinguir nada. Por último, las figuras E y F se presentan como ejemplo de falta de segregación (correspondería al caso de marcadores dominantes).

#### 4.2.2 Genotipado mediante SSR

Después de los mapas genéticos preliminares utilizando únicamente los SNPs, vimos que había 2 regiones del genoma de SD que eran heterocigotas y que no estaban cubiertas por marcadores. Debido a esto buscamos y genotipamos 2 marcadores SSR localizados en dichas regiones: CPPCT029 (1: 41.160.501) y CPPCT21395 (6: 1.130.915)<sup>15</sup>. Ambos marcadores son co-dominantes y se les consideró aptos para cubrir las dos regiones no cubiertas.

Para el genotipado mediante SSR se utilizaron las 418 muestras. El software entrega los resultados en forma de electroferograma<sup>16</sup>, dónde se muestran los valores alélicos de cada muestra. Los alelos se visualizan en forma de picos y se visualizan debido a la gran cantidad de copias que se han amplificado mediante PCR, y a que se han marcado con un fluorocromo, de color azul, en el caso del marcador CPPCT029 y verde en el marcador CPPCT21395. Si tenemos dos picos, para el mismo color, en la misma gráfica, significa que el individuo es heterocigoto y si sólo tenemos un pico, es que es homocigoto (Figura 12 y Figura 13)

La notación se transformó para poder trabajar en coherencia con los datos del genotipado SNP y SR, así pues, en un principio una región heterocigota se codificó como XY y una homocigota se codificó como XX, para el pico con mayor cantidad de pares de bases, y como YY para el de menor cantidad de pares de bases (Tabla 5).

Tabla 5 – Ejemplo del cambio de notación de los datos para poder trabajar con el mismo criterio y fusionarlos junto el genotipado mediante SNP. Los valores numéricos corresponden a los pares de bases (pb) dónde encontramos cada alelo.

Marcador	DATOS ORIGINALES		DATOS TRANSFORMADOS	
	CPPCT029 (pb)	CPPCT21395 (pb)	CPPCT029	CPPCT21395
SDF2-001	182/194	200	XY	XX
SDF2-002	194	187	XX	YY
SDF2-003	182/194	187/200	XY	XY
SDF2-004	194	187	XX	YY
SDF2-005	182/194	200	XY	XX
SDF2-006	182	187	YY	YY
SDF2-007	194	187/200	XX	XY
SDF2-008	182	187	YY	YY
SDF2-009	182/194	200	XY	XX
SDF2-010	182	187	YY	YY
SDF2-011	194	187/200	XX	XY
SDF2-012	182	187	YY	YY

<sup>15</sup> Nombre del marcador (grupo al que pertenece: posición física del cromosoma en pares de bases).

<sup>16</sup> Gráfico realizado con los resultados de un análisis por electroforesis. El electroferograma muestra la secuencia de datos producida por una máquina automática de secuenciación de ADN.

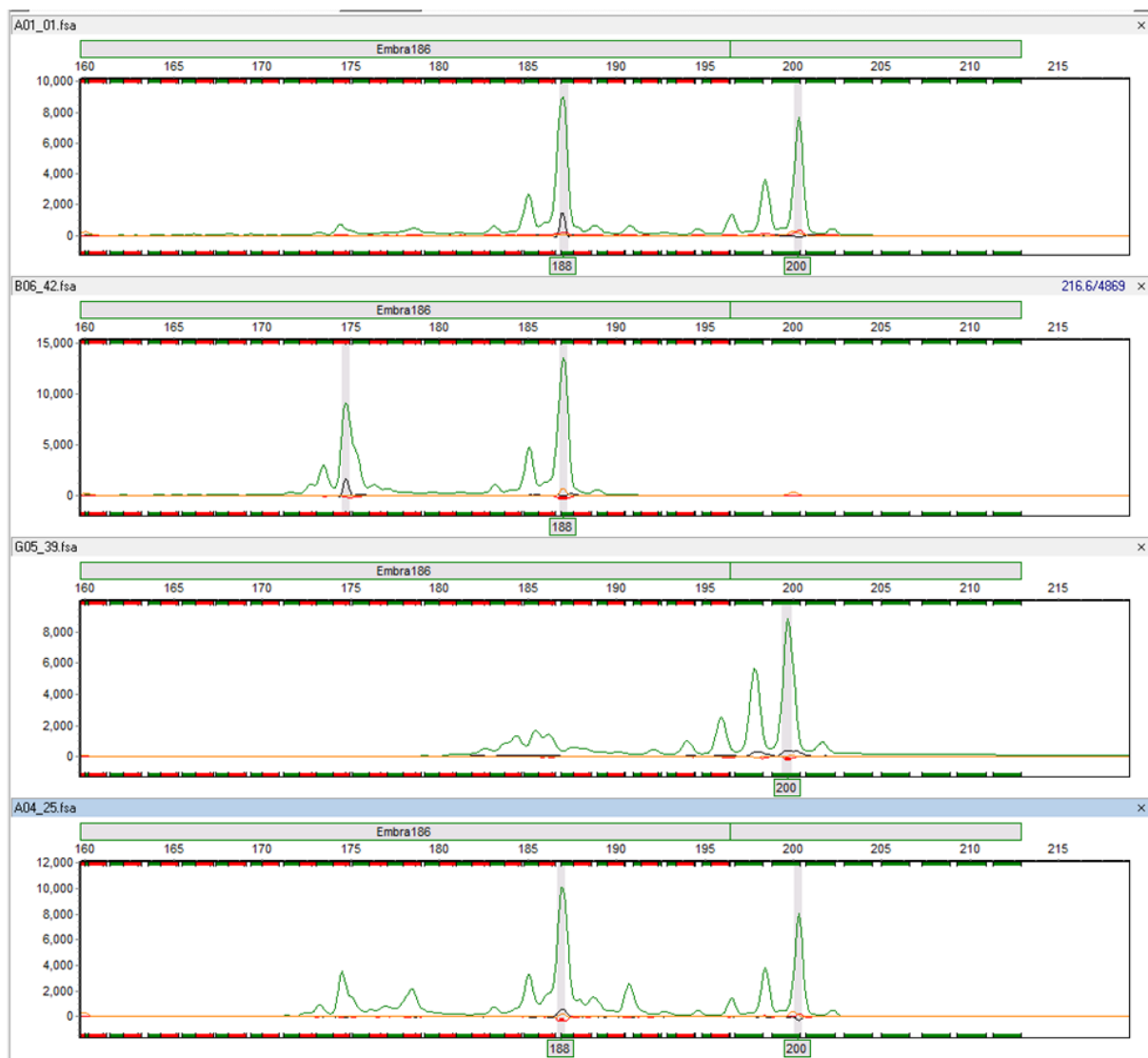


Figura 12 - Ejemplo de muestras analizadas mediante el programa GeneMarker (Softgenetics). En la figura se hace énfasis en los picos del marcador CPPCT21395 que entrega los resultados en pares de bases (187/200). La primera imagen corresponde al parental SD. Las 3 imágenes siguientes son individuos de la  $F_2$ . Si el individuo presenta ambos picos a la vez (187/200), como el caso de la primera y la última imagen, estamos ante un individuo heterocigoto, que se codificará como XY; si presenta el pico 187, como el caso de la segunda imagen, estamos ante un individuo homocigoto, que se codificará como YY y si presenta el pico 200, como el caso de la tercera imagen, estamos ante un individuo homocigoto, que se codificará como XX. En esta imagen se ha quitado el marcador de color azul para otorgar una perspectiva más visible del electroferograma.

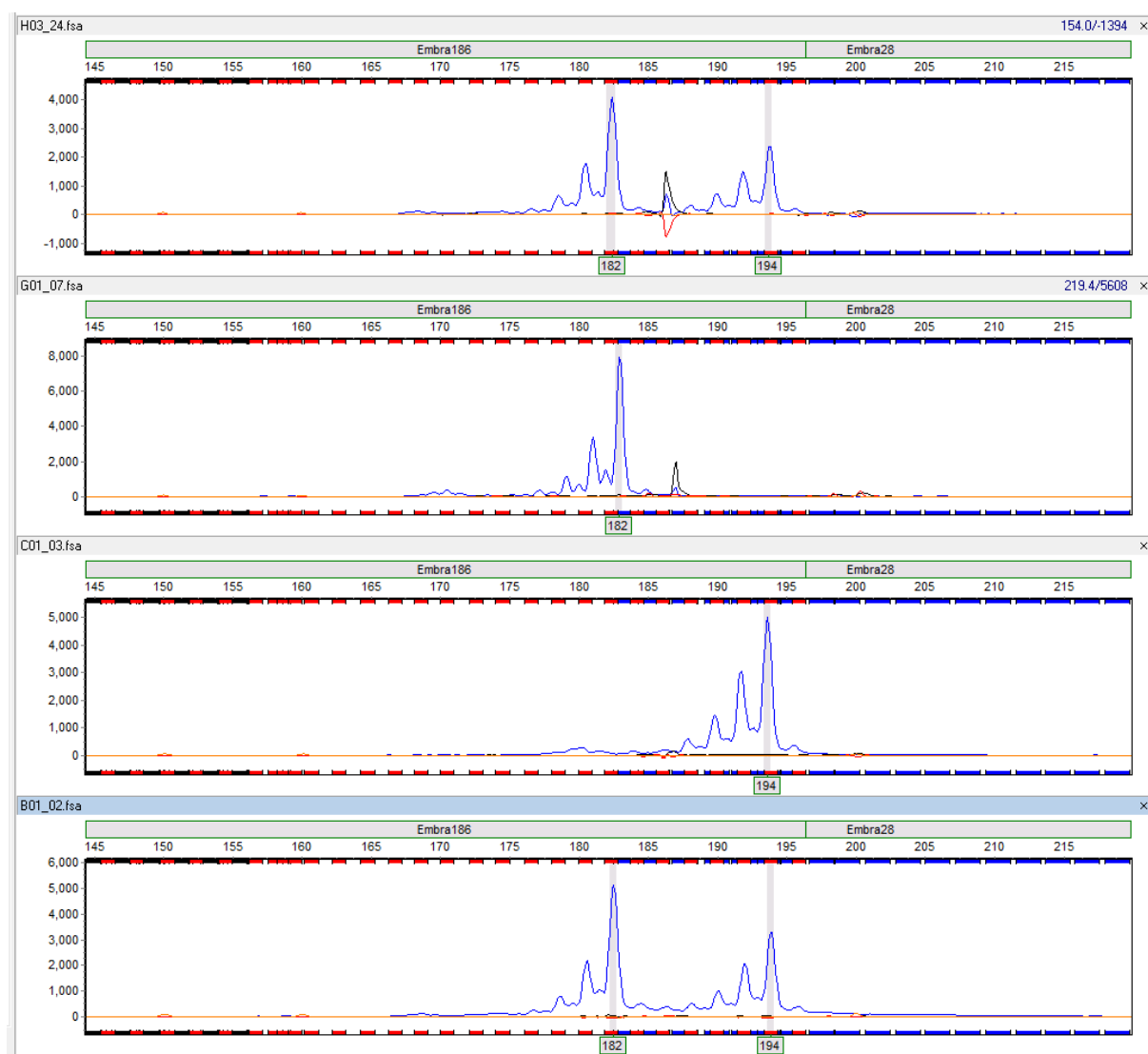


Figura 13 - Ejemplo de muestras analizadas mediante el programa GeneMarker (Softgenetics). En la figura se hace énfasis en los picos del marcador CPPCT029 que entrega los resultados en pares de bases (182/194). La primera imagen corresponde al parental SD. Las 3 imágenes siguientes son individuos de la F<sub>2</sub>. Si presenta ambos picos a la vez (182/194), como el caso de la primera y la última imagen, estamos ante un individuo heterocigoto, que se codificará como XY; si presenta el pico 182, como el caso de la segunda imagen, estamos ante un individuo homocigoto, que se codificará como YY y si presenta el pico 194, como el caso de la tercera imagen, estamos ante un individuo homocigoto, que se codificará como XX.

### 4.2.3 Genotipado slow ripening (SR)

De los 418 individuos disponibles se analizaron con el marcador SRcod un total de 60. Esto se debe a que solo se analizaron los individuos recombinantes entre los marcadores anterior y posterior al SRcod, es decir, si la región anterior y posterior son del tipo XX o XY será muy poco probable que presente el rasgo SR, ya que no es común encontrar una recombinación tan cerca de ambos marcadores. En la Tabla 6 se muestran 6 ejemplos que permiten aclarar el criterio descrito.



Tabla 6 – Ejemplo del criterio tomado para seleccionar los individuos que era necesario pasar el marcador co-dominante SR. El marcador SR está acotado por dos marcadores, uno que le antecede (SNP\_IGA\_409453) y otro que le precede (SNP\_IGA\_440116). Si en ambos marcadores se presenta XX y/o XY se sabe que estamos ante un individuo que difícilmente presentará el rasgo SR, por el contrario, si el individuo presenta YY/YY, YY/XY o YY/XX se sabe que estamos ante un individuo que puede presentar el rasgo SR y como no se sabe es necesario hacer un análisis para descartarlo o confirmarlo. En el individuo SDF2-201 es casi seguro que cuando pasemos el marcador SRcod nos de YY, para el individuo SDF2-048 es muy probable que si pasamos el marcador SRcod nos de YY, pero puede ser XY también y no ser un individuo SR. Para el resto de individuos es muy poco probable que se presente YY pues estaríamos delante de 2 recombinaciones en SDF2-400 y 192, y de 1 recombinación en SDF2-017 y 025, con lo que no se pasó el marcador SRcod.

Marcador	Crom.	Región	SDF2-201	SDF2-048	SDF2-400	SDF2-017	SDF2-025	SDF2-192
SNP_IGA_409453	Pp04	10.396.616	YY	YY	XX	XY	XX	XY
SRcod	Pp04	11.111.981 a 11.137.943	-	-	-	-	-	-
SNP_IGA_440116	Pp04	16.084.694	YY	XY	XY	XY	XX	XX
¿Puede ser un individuo SR?			SI (sr/sr)	SI (sr/sr)	NO (Sr/sr)	NO (Sr/sr)	NO (Sr/Sr)	NO (Sr/sr)

Se realizaron dos geles de agarosa, uno con las 15 primeras muestras más el parental Sweet dream (Figura 14) y otro gel con 45 muestras restantes.

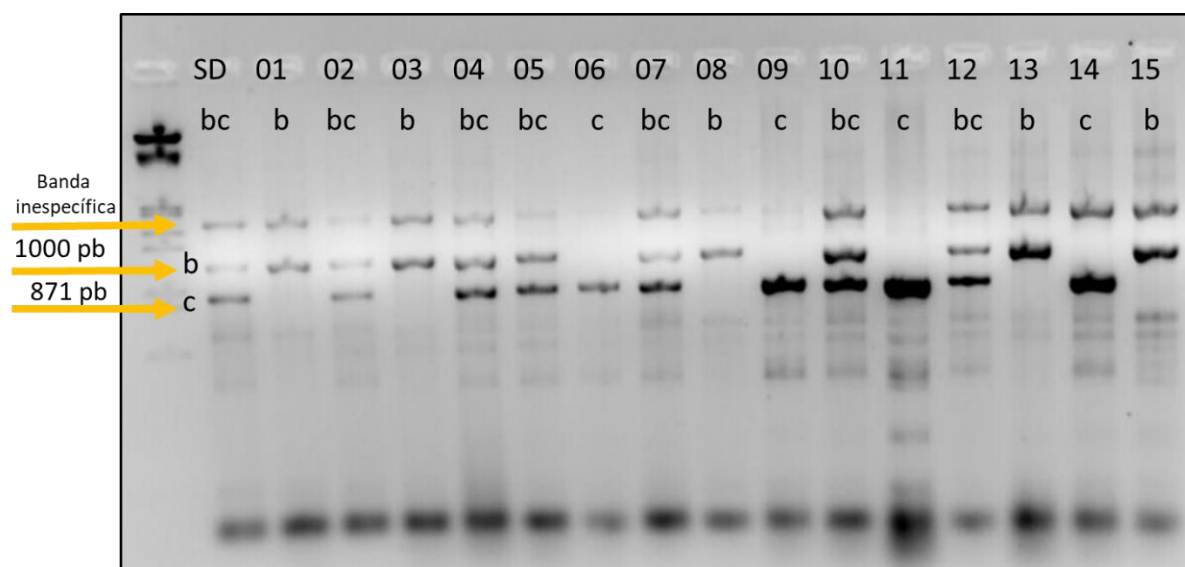


Figura 14 - Gel de agarosa de las muestras SDF2 de la 1 hasta la 15, más el parental Sweet Dream (SD), analizadas con el marcador SRcod. En el gel se identifican 3 bandas para los distintos tamaños de pares de bases: banda b, 1000 pb y banda c, 871 pb. Aquellos individuos que presentan sólo la banda b son individuos homocigotos que presentan los dos alelos Sr con lo que no segregan para el carácter SR, los individuos bc son individuos heterocigotos y presentan el alelo Sr y el sr, con lo que son portadores, y los individuos c son individuos homocigotos sr que les confiere el carácter SR. Los individuos SDF2-06,09 y 11 son SR (sr/sr), los individuos SDF2-02,04,05,07,10 y 12 son portadores (Sr/sr) y los individuos SDF2-01,08 y 13 no son portadores (Sr/Sr). Para tener todos los datos con la misma notación a los individuos b se les denominará XX, a los individuos bc XY y a los individuos c YY.

En total se detectaron 89 portadores individuos SR (Tabla 7).

**Tabla 7 - Individuos SR detectados en las 418 muestras analizadas.**

SDF2-006	SDF2-009	SDF2-011	SDF2-014	SDF2-021	SDF2-036	SDF2-045	SDF2-046	SDF2-047	SDF2-048
SDF2-050	SDF2-055	SDF2-056	SDF2-058	SDF2-059	SDF2-063	SDF2-064	SDF2-065	SDF2-072	SDF2-075
SDF2-083	SDF2-084	SDF2-086	SDF2-099	SDF2-102	SDF2-106	SDF2-122	SDF2-126	SDF2-128	SDF2-131
SDF2-132	SDF2-134	SDF2-135	SDF2-139	SDF2-142	SDF2-153	SDF2-155	SDF2-163	SDF2-167	SDF2-171
SDF2-185	SDF2-190	SDF2-191	SDF2-201	SDF2-204	SDF2-217	SDF2-227	SDF2-236	SDF2-237	SDF2-240
SDF2-245	SDF2-248	SDF2-253	SDF2-262	SDF2-265	SDF2-266	SDF2-268	SDF2-273	SDF2-274	SDF2-277
SDF2-289	SDF2-293	SDF2-296	SDF2-298	SDF2-300	SDF2-302	SDF2-307	SDF2-311	SDF2-312	SDF2-313
SDF2-319	SDF2-325	SDF2-331	SDF2-339	SDF2-345	SDF2-360	SDF2-368	SDF2-376	SDF2-378	SDF2-386
SDF2-389	SDF2-397	SDF2-399	SDF2-402	SDF2-403	SDF2-404	SDF2-405	SDF2-411	SDF2-419	

En la Figura 15 se presenta la diferencia entre un individuo SR i un individuo que no lo es. Se puede observar como el fruto de la izquierda ha llegado a un punto en el que no madurará más, por más tiempo que permanezca en el árbol.



**Figura 15 – Diferencia entre el fruto que da un individuo SR, imagen izquierda, con el fruto que da un árbol sin presentar el rasgo SR, imagen derecha.**

#### **4.2.4 Preselección de individuos**

Una vez finalizado el genotipado, mediante SNP, se procedió a eliminar aquellos individuos que presentaran más de un 10% de datos perdidos, con lo que se descartaron un total de 14 plantas (Tabla 8).

Tabla 8-Individuos descartados por presentar más de un 10% de datos perdidos.

INDIVIDUO	XX	XY	YY	Datos perdidos	% de datos perdidos
SDF2-020	9	35	11	9	13.6
SDF2-142	6	27	23	9	13.6
SDF2-158	21	19	15	9	13.6
SDF2-174	7	33	15	9	13.6
SDF2-166	32	13	9	10	15.2
SDF2-213	9	28	13	14	21.2
SDF2-146	18	7	24	15	22.7
SDF2-147	28	3	12	21	31.8
SDF2-144	23	5	14	22	33.3
SDF2-148	25	3	12	24	36.4
SDF2-149	21	1	18	24	36.4
SDF2-143	9	3	15	37	56.1
SDF2-145	17	0	9	38	57.6
SDF2-150	9	3	7	45	68.2

### 4.3 Mapa ligamiento genético

Con los datos de la matriz M1 actualizada (404 individuos y 66 marcadores), se elaboró el mapa de ligamiento, que sirvió para verificar que los marcadores utilizados proporcionan cobertura en todas las regiones heterocigotas, de todo el genoma, de la variedad Sweet Dream.

En la Figura 17 se muestra el mapa de ligamiento genético, de la población F<sub>2</sub>, de la variedad SD .

Cuando se trató de construir los mapas de ligamiento algunos de los marcadores utilizados no estaban en fase. Esto se debe a que no conocemos los alelos de los parentales de la variedad SD y se asignan los alelos al azar. Este cambio de fase se puede detectar gráficamente, ya que observamos a ambos extremos del cromosoma un aglutinamiento de los marcadores, fruto de las distancias exageradas entre los grupos; se observan distancias de centenares a miles de cM (Figura 16), cosa que no tiene sentido en un mapa de ligamiento, ya

que 1 cM equivale a 1% de frecuencia de recombinación<sup>17</sup> y la máxima frecuencia de recombinación es del 50%, lo que serían 50 cM en el mapa. Para solucionar este problema se toma el criterio de seleccionar todos los marcadores superiores o inferiores de todos los grupos y en la matriz M1 cambiar las XX por YY o viceversa, con lo que generamos el cambio de fase y ya se puede obtener el mapa de ligamiento completo.

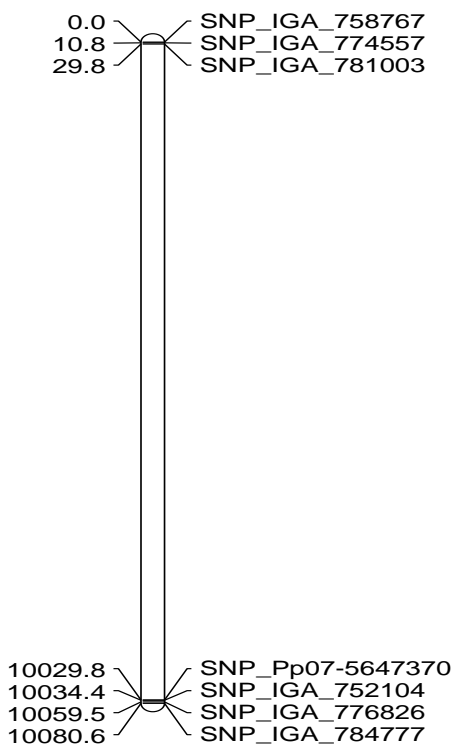


Figura 16 – Ejemplo del grupo 7 con marcadores cambiados de fase. Se puede observar cómo se aglutinan los grupos a cada extremo y la distancia entre ellos es de 10.000 cM

<sup>17</sup> La frecuencia de recombinación es un parámetro genético de cartografía genética, que, para dos loci dados, se emplea como indicador cuantitativo de la distancia que existe entre ambos. Su unidad, la unidad de mapa genético o centiMorgan (cM), como la frecuencia de recombinación de 0,01, esto es, del 1%, lo que quiere decir que uno de cada 100 productos de meiosis es recombinante.

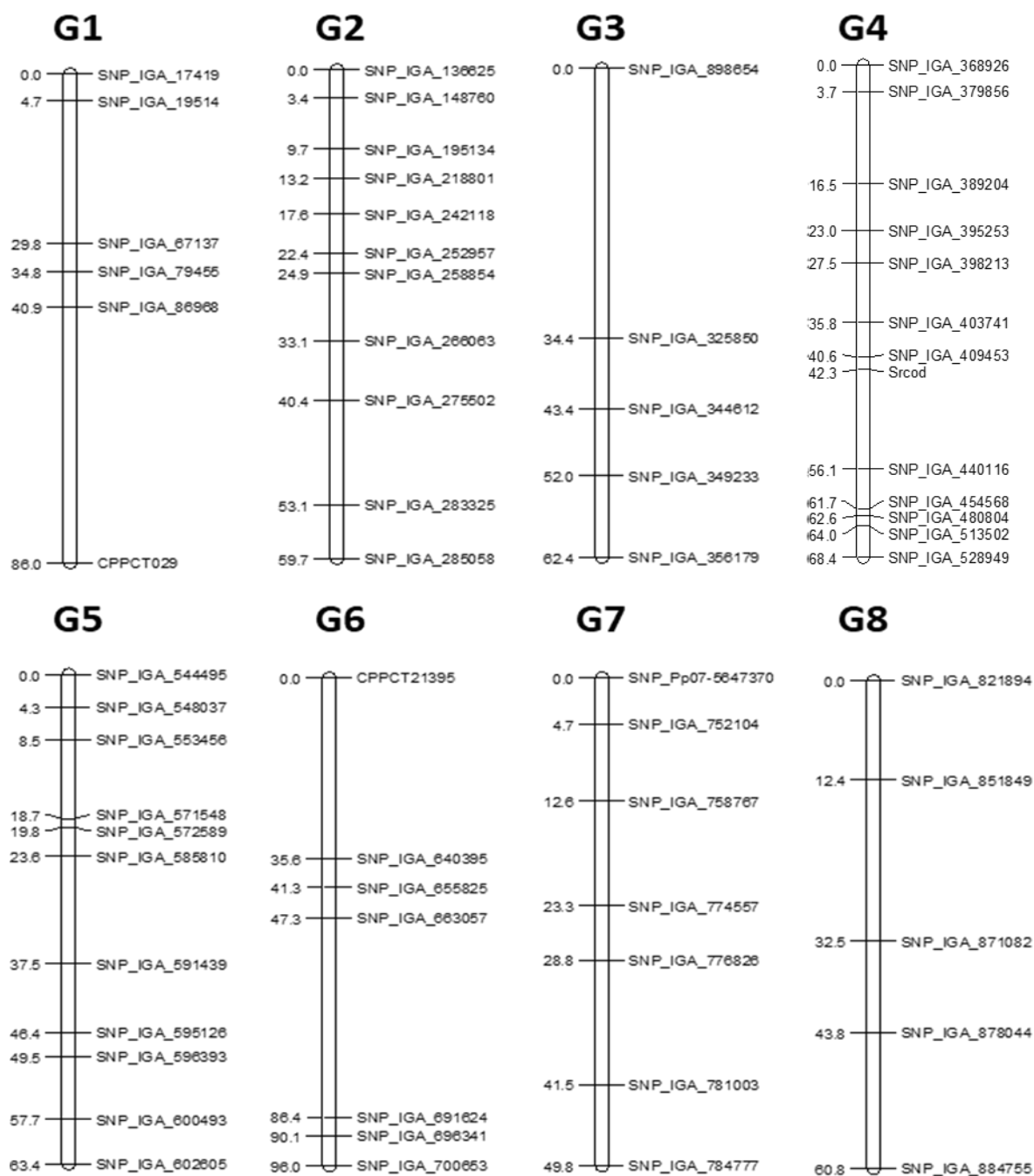


Figura 17 - Mapa de ligamiento genético de la población F<sub>2</sub> de la variedad Sweet Dream.

#### 4.4 Individuos aptos para la Resíntesis A

La Resíntesis A consiste en identificar, con pocos segmentos en homocigosis y a poder ser, cinco o más cromosomas totalmente heterocigotos, idénticos a los del cultivar original. Para lograr encontrar estos individuos, se escogieron aquellos que presentaban un valor de heterocigosidad observada  $\geq 0.75$ , reclutando un total de 11 individuos (Tabla 9).

Tabla 9 - Individuos seleccionados en función de la heterocigosidad observada.

Individuo	Heterocigosidad Observada
SDF2-039	0,92
SDF2-257	0,83
SDF2-374	0,82
SDF2-041	0,80
SDF2-017	0,78
SDF2-393	0,77
SDF2-276	0,77
SDF2-248	0,75 es SR
SDF2-258	0,75
SDF2-317	0,75
SDF2-136	0,75

De los individuos seleccionados por valor de heterocigosidad observada el individuo SDF2-248 presenta el rasgo SR.

Como la finalidad de este método es encontrar aquellos individuos que presenten la mayor cantidad de cromosomas heterocigotos totalmente idénticos al parental SD, se escogieron aquellos individuos con un valor de 4 o más cromosomas idénticos, obteniendo un total de 6 individuos (Tabla 10). Se observó que los individuos SDF2-285 y SDF2-168 obtuvieron valores de heterocigosidad observada, inferior al límite establecido inicialmente, pero presentaron 4 cromosomas idénticos al parental SD.

Tabla 10 - Individuos seleccionados en función de la cantidad de cromosomas en heterocigosis idénticos al parental (IHC).

Individuo	Heterocigosidad Observada	IHC	Cromosomas idénticos al parental Sweet Dream (SD)
SDF2-039	0,92	6	1, 3, 4, 5, 7, 8
SDF2-017	0,78	5	3, 4, 5, 6, 8
SDF2-374	0,82	4	1, 2, 4, 8
SDF2-257	0,83	4	4, 5, 6, 7
SDF2-285	0,69	4	1, 5, 6, 7
SDF2-168	0,57	4	2, 4, 6, 8

Del análisis de los 404 individuos se desprende que muy pocos individuos presenta 5 o más cromosomas completamente heterocigotos (CCH) (Tabla 11).

Tabla 11 – Cantidad de individuos ordenados en función de la cantidad de CCH.

CCH	Cantidad de individuos	% respecto a los 404 individuos analizados
0	140	34,7
1	139	34,4
2	88	21,8
3	31	7,7
4	4	1,0
5	1	0,2
6	1	0,2

#### 4.4.1 Individuos aptos para la Resíntesis B

##### 4.4.1.1 Correlación estadística

En el ANEXO 1 se entregan los resultados completos de la correlación estadística.

Una vez realizado el cálculo estadístico se observó que aquellos individuos que presentaron un mayor coeficiente de correlación negativo presentaron una mayor cantidad de regiones complementarias y una menor cantidad de regiones con 2 alelos homocigotos, a excepción del cruzamiento entre el individuo SDF2-131 con el SDF2-180 que presentó una sola región con 2 alelos homocigotos.

En la Tabla 13 se muestran los individuos más aptos para realizar un cruzamiento, que son los que presentaron mayor correlación estadística negativa.

Tabla 12 - Emparejamiento de individuos para realizar la Resíntesis B, en función de su coeficiente de correlación. Cuanto mayor sea el coeficiente de correlación negativo mayor será la cantidad de regiones complementarias y menor la cantidad de regiones con 2 alelos homocigotos.

Cruzamientos de individuos SDF2	Coeficiente de correlación
343 con 386	-0,80
037 con 086	-0,78
037 con 217	-0,75
028 con 292	-0,65
050 con 269	-0,6
131 con 180	-0,58
050 con 076	-0,54
238 con 037	-0,53
076 con 131	-0,53
180 con 207	-0,53

#### 4.4.1.2 Emparejamiento de individuos para la Resíntesis B

Para la resíntesis B se seleccionaron un total de 7 parejas de individuos que presentaron el menor valor de coeficiente de heterocigosidad ( $\leq 0.30$ ), juntamente con la mayor cantidad de regiones complementarias y regiones donde se combina heterocigoto con homocigoto, así como la menor cantidad de regiones con 2 alelos homocigotos (Figura 18). Los emparejamientos se ordenan de mayor a menor correlación estadística (Tabla 13).

Tabla 13-Emparejamientos de individuos aptos para realizar la técnica de la Resíntesis B. Nótese que presentan una gran cantidad de regiones complementarias y regiones dónde se da homocigoto más heterocigoto.

Parejas	343 vs 386	037 vs 086	037 vs 217	028 vs 292	050 vs 269	131 vs 180	050 vs 076	Leyenda
Correlación estadística	-0,8	-0,78	-0,75	-0,65	-0,6	-0,58	-0,54	
Región complementaria	39	37	35	33	35	26	34	
Homocigoto + heterocigoto	17	15	21	29	18	35	16	
2 alelos heterocigotos	5	9	7	1	3	0	6	
2 alelos homocigotos	1	2	0	1	8	1	8	
Missing Value	2	1	1	0	0	2	0	
Total	64	64	64	64	64	64	64	



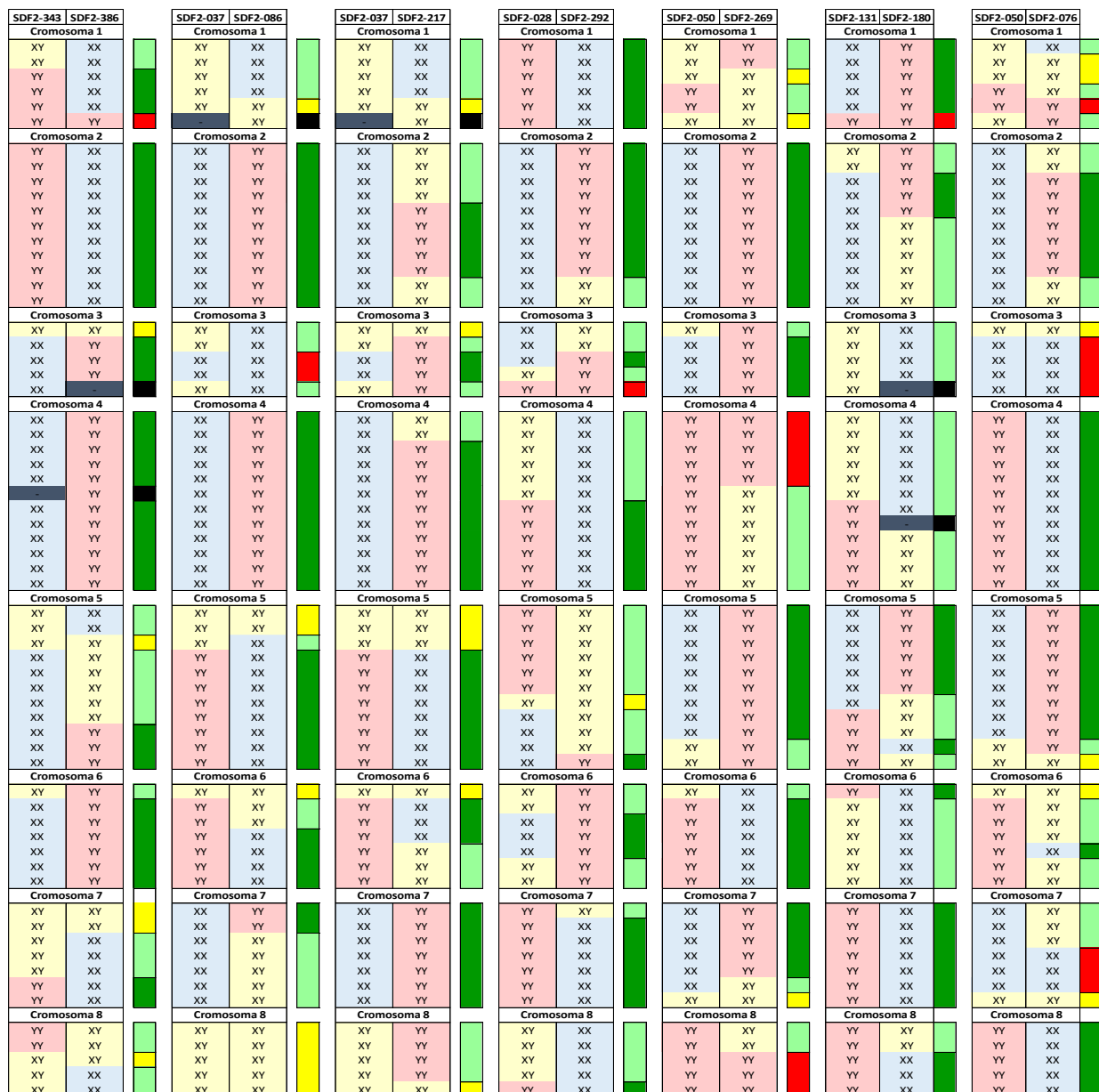


Figura 18- Los emparejamientos realizados presentan una columna de color en su margen derecho que representa el tipo de emparejamiento que tenemos entre los 2 individuos cruzados. Las regiones complementarias se marcan de color verde oscuro, las regiones donde se combina heterocigoto con homocigoto se marcan en color verde claro, las regiones con 2 alelos heterocigotos se marcan en color amarillo y las regiones con 2 alelos homocigotos se marcan en color rojo. Lo deseable es encontrar la mayor cantidad de regiones verde oscuro y verde claro, así como la menor cantidad de regiones de color rojo.

## 5 DISCUSIÓN

### 5.1 Genotipado mediante SNP

Mediante esta técnica se pudieron confirmar 64 marcadores con segregaciones mendelianas, aptos para la Resíntesis, de 96 analizados, lo que nos indica un 67% de éxito en la selección de marcadores.

Los marcadores SNPs, que no segregaron son aquellos en los que no es posible discernir entre homocigotos y heterocigotos.

El porcentaje de datos perdidos, en los 64 marcadores válidos, fue del 0.97%. Según el fabricante este tipo de placa tiene una precisión del 99%, con lo que estaríamos dentro del rango de datos perdidos admitidos. Realizar este análisis es necesario para saber si ha fallado algo en el análisis de las placas.

El uso de placas de matriz dinámica de circuito fluídico integrado, demostró ser una técnica muy útil, rápida y eficiente para el genotipado de SNPs. Permitió realizar 40.128 reacciones, en total, analizando 418 muestras con 96 marcadores, en tan solo 5 placas y en un tiempo total de 20 horas (4h por placa), lo que implica un máximo de 2-3 días de análisis. El mismo estudio mediante SSR habría requerido de la preparación de 120 placas, para poder ser analizadas en 60 sesiones, ya que el dispositivo analiza dos placas de 96 por sesión y teniendo en cuenta, que se pueden hacer 3 cargas diarias, habríamos destinado 20 días completos para el mismo análisis. También hay que tener en cuenta que el análisis de los SSR con programa informático es un proceso mucho más lento.

#### 5.1.1 Pérdida de datos (*missing values*)

La pérdida de datos se podría deber a tres factores:

1. Mala calidad del ADN utilizado.
2. Error introducido en el pipeteado de las muestras.
3. Los pocillos de los extremos de los chips son menos eficientes y en algunas ocasiones podemos tener una mayor cantidad de datos perdidos.

En nuestro caso se ha detectado un posible error en el pipeteado de las muestras SDF2-143 a SDF2-150, ya que la mayor cantidad de datos perdidos se da en estos individuos (Tabla 8) y están ubicados en la misma columna 2, de la placa 3, que se analizó mediante electroforesis capilar (Tabla 14). También se observó que 4 individuos con bastantes datos

perdidos se ubican en el extremo inferior de la placa y en este caso se puede deber a cualquiera de los tres factores. Para el caso del individuo SDF2-020 y SDF2-213 seguramente se debió a la mala calidad del ADN utilizado.

**Tabla 14 – Esquema de la placa 3 que se analizó mediante SNPs. Se puede observar como 13 de los 14 individuos con mayor cantidad de datos perdidos (>10%) se da en esta placa. Los individuos de la columna 2**

	1	2	3	4	5	6	7	8	9	10	11	12
A	SD	SDF2-143	SDF2-151	SDF2-159	SDF2-167	SDF2-175	SDF2-183	SDF2-191	SDF2-199	SDF2-207	SDF2-215	SDF2-223
B	SDF2-136	SDF2-144	SDF2-152	SDF2-160	SDF2-168	SDF2-176	SDF2-184	SDF2-192	SDF2-200	SDF2-208	SDF2-216	SDF2-224
C	SDF2-137	SDF2-145	SDF2-153	SDF2-161	SDF2-169	SDF2-177	SDF2-185	SDF2-193	SDF2-201	SDF2-209	SDF2-217	SDF2-225
D	SDF2-138	SDF2-146	SDF2-154	SDF2-162	SDF2-170	SDF2-178	SDF2-186	SDF2-194	SDF2-202	SDF2-210	SDF2-218	SDF2-226
E	SDF2-139	SDF2-147	SDF2-155	SDF2-163	SDF2-171	SDF2-179	SDF2-187	SDF2-195	SDF2-203	SDF2-211	SDF2-219	SDF2-227
F	SDF2-140	SDF2-148	SDF2-156	SDF2-164	SDF2-172	SDF2-180	SDF2-188	SDF2-196	SDF2-204	SDF2-212	SDF2-220	SDF2-228
G	SDF2-141	SDF2-149	SDF2-157	SDF2-165	SDF2-173	SDF2-181	SDF2-189	-	SDF2-205	SDF2-213	SDF2-221	SDF2-229
H	SDF2-142	SDF2-150	SDF2-158	SDF2-166	SDF2-174	SDF2-182	SDF2-190	SDF2-198	SDF2-206	SDF2-214	SDF2-222	-

## 5.2 Genotipado mediante SSR

Los dos marcadores SSR utilizados, para cubrir aquellas regiones que los marcadores SNP no pudieron cubrir, tuvieron un resultado satisfactorio y sirvieron para completar el final del grupo 1 y el principio del grupo 6.

## 5.3 Genotipado slow ripening (SR)

El carácter *slow ripening* está controlado por un único gen (Sr/sr), que, Como es la forma recesiva homocigota (sr/sr) la que confiere el rasgo SR, la cantidad de individuos esperados que presenten este fenotipo será del 25%.

Siguiendo la metodología de Meneses et al (2016) se pudieron detectar 89 individuos portadores del fenotipo SR, lo que indica que el 22% de los individuos analizados serían SR. Resultado que no se aleja mucho del teórico y por lo tanto esperable.

La teoría nos indica la cantidad esperable de individuos SR, pero no nos indica cuales van a ser. Con los marcadores del grupo 4, SNP\_IGA\_409453 y SNP\_IGA\_440116 se pudo observar que individuos eran susceptibles de ser SR, pero en los casos que no se podía identificar mediante los marcadores SNP se tuvo que recurrir al marcador SRcod que los identificó de forma directa.

Nos interesa descartar aquellos individuos que son SR para la Resíntesis A porque son los individuos que van a ser fenotipados i comparados con el cultivar original SD, es decir, el

resultado de la Resíntesis A son individuos aptos para utilizar. En cambio, para la Resíntesis B no hace falta descartar los individuos SR porque la finalidad son sus descendientes,. Como aclaración a este último punto; como estamos cruzando individuos con regiones homocigotas complementarias provenientes de la autofecundación de SD, cuando se cruce un individuo srsr con un SrSr tendremos un individuo Srsr, será portador, pero no será un individuo SR con lo que no se descartará para la Resíntesis B aquellos individuos que sean SR.

## **5.4 Mapa de ligamiento genético**

El mapa de ligamento se presentó como una herramienta indispensable para detectar que regiones del genoma del SD quedaban sin cubrir mediante el genotipado con SNPs. Una vez identificadas aquellas regiones sin cobertura se utilizaron marcadores del tipo SSR para volver a realizar el mapa de ligamiento y ver si efectivamente cubrían las regiones dónde nos faltaban marcadores.

## **5.5 Resíntesis**

Según Arús, et al (2016) si se ensayaran unos 2000 individuos se podrían obtener, a nivel teórico:

- 50-60 individuos con 5 o más cromosomas completamente heterocigotos (CCH), para la Resíntesis A.
- Individuos con 5 o más cromosomas completamente homocigotos (CCHm) y parejas de individuos con 4 o más cromosomas totalmente complementarios, para la Resíntesis B.

Para el estudio de la Resíntesis se trabajó con un total de 404 individuos cuando lo ideal debería haber sido trabajar con una población de unos 2000 individuos tal como se indica en Arús, et al (2016), tanto para el caso de la Resíntesis A como B.

### **5.5.1 Resíntesis A**

Los individuos SDF2-017 y SDF2-039 presentaron 6 CCH y 5 CCH, respectivamente. Esto representa un 0.5% de los individuos ensayados en comparación del 3% que estiman los cálculos teóricos.

En la Tabla 11 se puede ver los resultados dónde se observa que el 99.5% de individuos presenta  $\leq 4$  CCH.

Se ha constatado que el hecho de tener un coeficiente de heterocigosidad mayor no implica tener una mayor cantidad de CCH, tal y como esperábamos ya que el número de CCH depende de las recombinaciones que se dan y que son al azar. En la Tabla 15 se puede observar como individuos con valores de heterocigosidad  $\geq 0.70$  presentan 2 o 1 CCH.

Tabla 15 – Individuos con valores altos de heterocigosidad observada pero con  $\leq 2$  CCH.

Heterocigosidad Observada	Individuo	CCH
0,80	SDF2-041	2
0,75	SDF2-136	2
0,75	SDF2-276	2
0,75	SDF2-317	2
0,75	SDF2-248	1
0,73	SDF2-208	2
0,73	SDF2-400	1
0,72	SDF2-308	2
0,72	SDF2-356	2
0,72	SDF2-015	1
0,72	SDF2-074	1
0,71	SDF2-204	2
0,70	SDF2-002	2
0,70	SDF2-179	2

Existen varios individuos con cromosomas que son totalmente heterocigotos pero que no se puede asegurar que lo son, ya que presentan un dato perdido en uno de los marcadores del grupo, si se optara por suponer que ese dato perdido se corresponde a una región heterocigota se podrían añadir a la lista de individuos aptos para la Resíntesis A, estos son: SDF2374, SDF2-393, SDF2-412 y SDF2-413. Con lo que se podría obtener 6 individuos aptos de la Resíntesis A.

El problema que nos hemos encontrado es que se debería haber trabajado con una mayor cantidad de individuos y si la variedad SD no lo permite se debería buscar otro cultivar que lo permita.

Los individuos seleccionados de la Resíntesis A se deberán fenotipar y encontrar aquellos individuos que sean lo más similares al cultivar original SD pero con las características que nos interesen incluir, como por ejemplo ver aquellos que maduren antes o más tarde, con lo que podríamos extender el período comercial del cultivar original.

### 5.5.2 Res ntesis B

El individuo SDF2-238 present  5 CCHm y coeficiente de heterocigosidad de 0. Pero no se ha encontrado ninguna pareja de individuos con 4 o m s cromosomas totalmente complementarios. Tan s lo la pareja SDF2-037 con SDF2-084 tienen 3 cromosomas totalmente complementarios.

Se ha constatado que el hecho de tener un coeficiente de heterocigosidad muy bajo est  relacionado con tener una mayor cantidad de CCHm. En principio se espera que a menor coeficiente mayor cantidad de CCHm, pero no existe una relaci n lineal. En la Tabla 16 se puede observar como individuos con valores de heterocigosidad  $\geq 0.20$  presentan 4 CCHm e individuos con valores de heterocigosidad  $\leq 0.20$  presentan 3 CCHm, incluso individuos con un valor de heterocigosidad  $> 0.50$  pueden presentar 3 CCHm.

Tabla 16 – Individuos ordenados por la mayor cantidad de cromosomas completamente homocigotos. En la tercera columna se muestran la cantidad de cromosomas homocigotos con dos alelos X, en la cuarta columna se muestran la cantidad de cromosomas homocigotos con dos alelos Y y en la  ltima se entrega la suma total de cromosomas completamente homocigotos.

Individuos	Heterocigosidad	2 alelos homocigotos XX	2 alelos homocigotos YY	Total XX + YY
SDF2-238	0	3	2	5
SDF2-269	0,23	1	3	4
SDF2-292	0,23	3	1	4
SDF2-399	0	1	2	3
SDF2-050	0,14	1	2	3
SDF2-060	0,19	2	1	3
SDF2-386	0,19	1	2	3
SDF2-028	0,25	1	2	3
SDF2-180	0,26	2	1	3
SDF2-086	0,27	1	2	3
SDF2-037	0,27	3	0	3
SDF2-207	0,29	2	1	3
SDF2-225	0,29	1	2	3
SDF2-227	0,30	2	1	3
SDF2-130	0,31	1	2	3
SDF2-156	0,32	0	3	3
SDF2-084	0,33	0	3	3
SDF2-093	0,33	2	1	3
SDF2-157	0,37	1	2	3
SDF2-188	0,37	2	1	3
SDF2-322	0,40	1	2	3
SDF2-089	0,41	1	2	3
SDF2-135	0,41	1	2	3
SDF2-167	0,41	2	1	3
SDF2-404	0,41	1	2	3
SDF2-032	0,44	1	2	3
SDF2-057	0,52	0	3	3
SDF2-103	0,53	2	1	3

Encontrar emparejamientos posibles con 4 o más cromosomas totalmente complementarios en nuestro ensayo es totalmente imposible, ya que 141 individuos tienen 0 CCHm, 163 tienen 1 CCHm, 72 tienen 2 CCHm y 25 tienen 3 CCHm, con lo que representan el 99.25% del total y al emparejarse entre ellos en el mejor de los casos se obtendrían 3 cromosomas completamente complementarios.

Los 3 individuos con 4 o más CCHm si se emparejan entre ellos presentan cantidades inasumibles de regiones con 2 alelos homocigotos, los tres emparejamientos posibles producirán:

23 regiones de 64, con 2 alelos homocigotos para SDF2-238 con SDF2-269.

35 regiones de 64, con 2 alelos homocigotos para SDF2-238 con SDF2-292.

13 regiones de 64, con 2 alelos homocigotos para SDF2-269 con SDF2-292.

El cálculo estadístico para encontrar emparejamientos de individuos ha resultado satisfactorio y ha permitido encontrar los mejores emparejamientos posibles sin tener que realizar la cruce de forma manual. En la Figura 19 se puede observar como el emparejamiento con correlación estadística negativa (-0.78) presenta una gran cantidad de regiones complementarias y sólo 2 regiones con 2 alelos homocigotos, en cambio el emparejamiento con correlación estadística de 0 muestra pocas regiones complementarias y la correlación estadística positiva (0.75) presenta sólo 1 región complementaria y 38 regiones con 2 alelos homocigotos, con lo que escogiendo aquellos individuos con correlación negativa obtendremos las mejores parejas para obtener la mayor cantidad de regiones complementarias.

[illegible]

Figura 19 – Ejemplo de correlación estadística entre individuos con correlación negativa, 0 y positiva. Las regiones complementarias se marcan de color verde oscuro, las regiones donde se combina heterocigoto con homocigoto se marcan en color verde claro, las regiones con 2 alelos heterocigotos se marcan en color amarillo y las regiones con 2 alelos homocigotos se marcan en color rojo. Lo deseable es encontrar la mayor cantidad de regiones verde oscuro y verde claro, así como la menor cantidad de regiones de color rojo.



Con las parejas de individuos seleccionadas para la Resíntesis B se debería realizar el cruzamiento entre ellas y fenotipar los individuos seleccionados para compararlos con el cultivar original. En el caso de que se obtuvieran parejas de individuos con muchas regiones con los 2 alelos heterocigotos (color amarillo) o con 1 alelo homocigoto con otro heterocigoto (verde claro), se debería realizar la selección asistida por marcadores (SAM) para tratar de encontrar aquellos individuos con la mayor cantidad de regiones complementarias (verde oscuro).

## **6 IMPACTO ECONÓMICO DE LOS RESULTADOS**

### **6.1 Reducción de los tiempos y costes de selección varietal mediante el uso de marcadores moleculares**

Hasta finales del siglo XX las mejoras varietales del melocotón se han realizado mediante mejora genética clásica, lo que consiste en: recolectar el polen, realizar hibridaciones en campo, esperar a que el árbol de frutos y recolectarlos para cultivar embriones y/o germinar semillas, para posteriormente ser trasplantadas en campo y seleccionar los individuos en base a su fenotipo. Este proceso conlleva evaluar cada árbol, y mantenerlos por lo menos durante 4 años en el campo experimental. Además, si lo que se pretende es registrar una variedad, puede conllevar una media de 10 años (Pérez, 2012).

El problema de la mejora genética clásica es que se necesitan invertir recursos para mantener los cultivos en fase experimental (agua, fertilizantes, productos fitosanitarios, mano de obra, combustible, electricidad...) lo que conlleva un coste económico y ambiental.

Los marcadores moleculares permiten seleccionar aquellos individuos con las características deseadas, de la variedad objetivo, sin tener que esperar al análisis de los frutos. El proceso tan sólo se requiere germinar las semillas, plantar en semillero, dejar que broten las primeras hojas y extraer el ADN de las más jóvenes para hacer el genotipado. Con los datos obtenidos se sabrá de antemano que individuos nos interesan y cuáles no en tan solo unos meses.

#### **6.1.1 Caso de estudio de este trabajo**

Sírvase como ejemplo el caso analizado en este trabajo, en el que se ha seleccionado un cultivar de interés económico como la variedad Sweet Dream, con la que se ha realizado una autofecundación, obteniendo una población F<sub>2</sub>. De la población resultante se han analizado para ver que individuos serán portadores de una mutación, denominada maduración lenta (*slow rippening*), mediante un genotipado con los marcadores SNPs y el marcador SRcod. Si no se hubiera aplicado esta técnica deberíamos haber plantado todos los individuos en campo y esperar a observar los primeros frutos para saber que árboles son SR.

En nuestro caso se estuvo recolectando semillas durante 3 años (2015,2016,2017) para su posterior germinación, con la finalidad de obtener el material vegetal suficiente y hacer el genotipado en tan solo unas semanas. Si este proceso se hubiera realizado de forma

tradicional habríamos tardado 5-6 años, ya que deberíamos haber plantado los individuos y habríamos obtenido los resultados en 2019,2020 y 2021. En cambio los resultados se han obtenido en 2018, lo que supone una reducción del 50% en tiempo.

No obstante, si nuestra variedad hubiese tenido una tasa de germinación más alta con los 2000 individuos del 2015 quizás ya habría sido suficiente y en tan sólo 1 año podríamos haber llevado acabo el análisis.

## 6.1.2 Ejemplo real de un estudio de mejora genética

### 6.1.2.1 Costes asociados a mantener un cultivo de melocotonero

Los costes que se entregan en este apartado son de carácter general, ya que los costes dependen de muchos factores, tales como la variedad escogida, el marco de cultivo que nos dará la cantidad de árboles por hectárea, las condiciones climatológicas, el tipo de suelo dónde se realiza el cultivo, la cantidad de hectáreas cultivadas (Figura 20), etc.

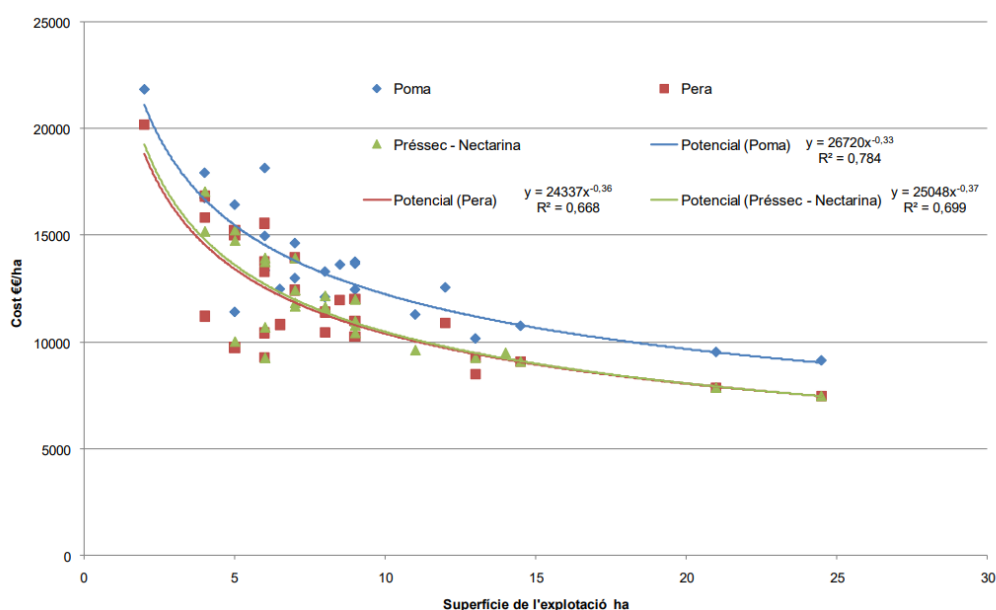


Figura 20 – Costes en €/ha en función de la superficie de explotación de diversos cultivos. La línea con triángulos verdes corresponde al cultivo del melocotón. En el gráfico se puede observar que pasar de una explotación de 1 a 10 hectáreas supone una reducción de costes del 50%. Fuente: (Observatori de la fruita fresca, 2018).

En la Tabla 17 se entrega un desglose de los costes asociados, en euros por hectárea (€/ha), a mantener un cultivo de melocotón. Según estos datos se desprende que mantener una hectárea de melocotón cuesta alrededor de los 6.500 €/ha. Para no tener un único dato se consultaron los datos del Observatori de la fruita fresca (2018), que según sus cálculos mantener una explotación de menos de 8 ha rondaría 8.280 €/ha y una explotación de más de 8 ha rondaría los 7.756 €/ha.

Tabla 17 - Desglose de los costes asociados, en euros por hectárea (€/ha), a mantener un cultivo de melocotón por año.  
Fuente: <http://www.paeria.es/fitxers/tauler/2008/08-097GECR/MEMORIES/6-VAL/MEMORIES/2-ANX/Annex4.pdf>

	Horas m.o./ ha	Horas maquin/ha	Coste de la m.o €/ha	Coste de los consumos €/ha	Gasoli- Repar Manten y seguro €/ha	Total costos €/ha
Tratamientos fitosanitarios varios	15,00	15,00	499,75	781,36	344,14	1.625,25 €
Tratamientos herbicida	2,77	2,77	157,66	46,88	63,55	268,09 €
Repartir abono	0,87	0,87	45,87	84,82	19,88	150,56 €
Repartir quelaots	1,70	1,70	56,64	262,45	39,00	358,09 €
Regar	2,00	0,00	9,60		0,00	9,60 €
Podar	68,00	10,00	379,69		18,53	398,21 €
Raecoger ramilla	1,40	1,40	48,48		32,12	80,60 €
Desbrozar ramas y hierba	10,00	10,00	393,58		229,43	623,01 €
Aclareo de fruto	25,00	0,00	120,00	0,00	0,00	120,00 €
Otros 1	0,00	0,00	0,00		0,00	- €
Otros 2	0,00	0,00	0,00		0,00	- €
Trabajos sin maquinas	0,00	0,00	0,00			- €
<b>Total cose producción excluido cosechar</b>	<b>126,74</b>	<b>126,74</b>	<b>1711,27</b>	<b>1175,50</b>	<b>746,64</b>	<b>3633,41</b>
Cosechar	190,00	85,00	1364,92		157,47	1.522,40 €
Transporte fruta dentro finca	0,44	0,44	113,00		10,09	123,09 €
Transporte fruta al almacen	27,53	21,00	224,03		481,79	705,82 €
<b>T.costos directos cultivo</b>	<b>344,27</b>	<b>232,74</b>	<b>3.300,22</b>	<b>1.175,50</b>	<b>1.385,91</b>	<b>5.861,63 €</b>
Coste capital circulante(costos externos)						<b>73,27 €</b>
<b>Total costos cultivo €/ha</b>						<b>5.934,90 €</b>
Amortitzación + coste capital plantaciones						162,32
Seguro de riesgo						72,77
Canon agua						88,32
Cotizaciones a la Seguridad Social						150,62
Energía eléctrica						50,00
Teléfono						0,00
Otros gastos						52,47
Coste capital gastos de estructura						5,18
<b>Total gastos de estructura por ha</b>						<b>576,50</b>
<b>Costos totales por ha</b>						<b>6.511,40 €</b>

### 6.1.2.2 Estudio realizado por IRTA-ASF

Se utiliza como ejemplo el trabajo realizado por el IRTA, junto con ASF (*Agro Selection Fruits*) dónde se han conseguido un total de 14 nuevas variedades de melocotón<sup>18</sup> y nectarina que han sido registradas en Europa (CPVO, Community Plant Variety Office): 2 melocotones de carne amarilla; 2 melocotones de carne blanca; 4 nectarinas de carne amarilla; 5 nectarinas de carne blanca; y 1 nectarina plana de carne blanca.

<sup>18</sup><https://www.lavanguardia.com/local/lleida/20130729/54378963787/presentan-variedades-catalanas-melocoton-nectarina.html>

El proyecto se inició en 2004 y finalizó en 2012. Se tardaron 8 años para obtener estas variedades. Para ello se realizaron 935 cruces, se obtuvieron 71.367 semillas y se evaluaron un total de 24.170 árboles, de los que se seleccionaron finalmente 318 para obtener las 14 variedades.

Los costes estimados de mantener todos los árboles durante los 5-6 años que estarían en desarrollo podrían haber representado un coste total de: 1.3 a 1.5 millones de euros.

Los cálculos se realizan en base a la información que proporciona Batlle, et al (2012), dónde dice que se utilizó un marco de plantación de 5x3m, con lo que se podrían cultivar unos 667 árboles/ha (cálculo estimado). Si tenemos en cuenta que se tendrían que mantener 24.170 árboles, esto supone un espacio de 36 ha, que poniendo un coste medio de 7.200€/ha anual, y que los árboles se estarán en campo 5-6 años. Se toma como referencia 5-6 años; aunque el ensayo duró 8, ya que los árboles se tuvieron que obtener de la germinación de semillas y las plántulas no generan un gasto tan grande. Además, hay que tener en cuenta que los datos de coste €/ha hacen referencia a campos en producción y estos son árboles completamente desarrollados.

Si tenemos en cuenta los datos de la Tabla 18, el coste del fenotipado por individuos (trasplante, mantenimiento en finca 3 años, evaluación estadio 1, eliminación del árbol) cuesta unos 6.5-7.5€/individuo como máximo, con lo que si se analizaran los 24.170 individuos tendríamos unos costes de 157.105 a 181.275 €.

De los cálculos estimativos se desprendería que utilizar SAM supone una reducción de costes de entre 8 a 10 veces menos que con los métodos de la mejora convencional. Además, el hecho de no utilizar tantos recursos genera un impacto mucho menor en el medio ambiente, que no es del alcance de este trabajo.

Sabemos que en 2008 se inició una colaboración de los mejoradores del programa IRTA-ASF-Fruit-Futur con investigadores del Programa de Genómica y Biotecnología del IRTA situado en el CRAG (*Centre de Recerca Agrigenòmica*). Se utilizaron 48 marcadores SSR para proporcionar una buena cobertura del genoma del melocotón. Se identificaron caracteres para los que podría ser interesante para realizar selección asistida por marcadores. Esto representó un ahorro importante en costes. Pero se quería hacer una estimación de que hubiera pasado si se hubiera llevado a cabo todo el programa sin la herramienta de la SAM.

**Tabla 18 - Comparación de costes entre selección convencional y SAM. Los costes calculados incluyen material y mano de obra desde el estado de plántula en invernadero hasta la primera toma de decisión de eliminación/selección en la parcela de descendencias. Fuente: (Eduardo, et al., 2015).**

<b>Objetivo</b>	<b>Nº individuos a genotipar</b>	<b>Nº individuos descartados</b>	<b>Precio genotipado<sup>1</sup> + fenotipado individuos seleccionados (€)</b>	<b>Precio fenotipado<sup>1</sup> (€)</b>
Descartar individuos homocigotos de un carácter recesivo indeseado donde ambos padres son portadores	1.000	250	7.625–9.125	6.500–7.500
Seleccionar individuos portadores de un carácter dominante deseado donde un padre es portador	1.000	500	6.000–7.250	6.500–7.500
Seleccionar individuos homocigotos de un carácter recesivo deseado donde ambos padres son portadores	1.000	750	4.375–5.375	6.500–7.500
Seleccionar individuos portadores para dos caracteres dominantes	1.000	750	5.125–6.125	6.500–7.500

(1) 2,75–3,50€ genotipado x individuo (extracción + 1 SSR); 3,50–4,25€ (extracción + 2SSRs).

(2) 6,5–7,5€ fenotipado x individuos (trasplante, mantenimiento en finca 3 años, evaluación estadio 1, eliminación del árbol).

## 7 CONCLUSIONES

De los 96 SNPs analizados para la Resíntesis 64 se pudieron utilizar. El porcentaje de datos perdidos, en estos 64 marcadores fue del 0.97%.

El mapa de ligamento se presentó como una herramienta indispensable para detectar que regiones del genoma del SD quedaban sin cubrir mediante el genotipado con SNPs.

Los dos marcadores SSR utilizados tuvieron un resultado satisfactorio y sirvieron para completar el final del grupo 1 y el principio del grupo 6.

Para la Resíntesis A se han seleccionado 6 individuos en base a la cantidad de regiones completamente homocigotas. Los únicos individuos con 5 o más CCH fueron el SDF2-017 y el SDF2-039. Esto representa un 0.5% de los individuos ensayados en comparación del 3% que estiman los cálculos teóricos.

Los individuos seleccionados de la Resíntesis A se deberán fenotipar para ver si encontramos individuos similares a SD pero con una alguna característica diferencial e interesante.

Para la Resíntesis B, el cálculo estadístico de las correlaciones para encontrar emparejamientos de individuos ha resultado satisfactorio y ha permitido encontrar los mejores emparejamientos posibles. A pesar de no haber encontrado ninguna línea con más de 6 CCHm, se han seleccionado 7 parejas complementarias en base a CCHm y el mayor coeficiente de correlación negativo.

Se han identificado 3 individuos (SDF2-238, SDF2-269 y SDF2-292) con 4 o más cromosomas completamente homocigotos, pero no sus parejas complementarias.

Con las parejas de individuos seleccionadas para la Resíntesis B se debería realizar el cruzamiento entre ellas, seleccionar los individuos más similares a SD utilizando SAM y fenotipar los individuos seleccionados para compararlos con el cultivar original.

Con los resultados obtenidos, y a falta de los fenotipos, podríamos concluir que la estrategia de la Resíntesis es una estrategia factible, incluso con una variedad muy heterocigota y analizando menos individuos de los que se deberían en base al cálculo teórico.

Se pudieron detectar 89 individuos portadores del fenotipo SR, lo que indica que el 22% de los individuos analizados serían SR.

## 8 BIBLIOGRAFÍA

Albillos García, S. M., 2003. *Aplicación de la electroforesis capilar para el análisis y seguimiento del grado de maduración de quesos de oveja y mezcla de la provincia de Burgos. Tesis doctoral inédita*.. Burgos: Universidad de Burgos. Facultad de Ciencias..

Aranzana MJ, E. K. A. H. W. A. P., 2010. Genetic variation, population structure and linkage disequilibrium in peach commercial varieties.. *BMC Genetics*, Issue 11, p. 69.

Aranzana, M., García-Mas, J., Carbó, J. & Arús, P., 2002. Development and variability analysis of microsatellite markers in peach. *Plant Breeding*, 121(1), pp. 87-92.

Arús, P., Howard, W., Aranza, M. J. & Eduardo, I., 2016. *RESYNTHESIS and VORIs: Two new marker-based approaches for peach breeding. PLANT & ANIMAL GENOME XXIV (PAG XXIV). Conference*. San Diego, California, USA, s.n.

Batlle, I. y otros, 2012. *The peach breeding programme IRTA-ASF: Aiming for high fruit quality. Acta horticulturae 940*, 75-78. Lisbon (Portugal), s.n.

Benavides , F. J. & Guénet, J.-L., 2003. *Manual de genética de roedores de laboratorio. Principios básicos y aplicaciones*.. Berlin: Universidad de Alcalá de Henares y la SECAL.

Cooperativas Agro-alimentarias , 2018. *Cooperativas Agro-alimentarias*. [En línea] Available at: <http://www.agro-alimentarias.coop/ficheros/doc/05696.pdf> [Último acceso: 10 Agosto 2018].

Cornejo Romero, A., Serrato Díaz, A., Rendón Aguilar, B. & Rocha Munive, M. G., 2014. *Herramientas moleculares aplicadas en ecología: aspectos teóricos y prácticos*. Ciudad de México (DF): Instituto Nacional de Ecología y Cambio Climático (INECC).

Dagnino, J., 2014. Bioestadística y epidemiología. Datos faltantes (missing values). *Revista chilena de anestesia*, 43(4), pp. 332-334.

Doyle, J. & Doyle, J., 1990. Isolation of Plant DNA from Fresh Tissue. *Focus*, 12(1), pp. 13-15.

Eduardo, I., Cantín, C., Batlle, I. & Arús, P., 2015. Integración de los marcadores moleculares en un programa de mejora de variedades de melocotonero. *Fruticultura*, Issue 44, pp. 6-17.



Eduardo, I. y otros, 2015. Mapping of a major gene for the slow ripening character in peach: co-location with the maturity date gene and development of a candidate gene-based diagnostic marker for its selection.. *Euphytica*, Volumen 205, pp. 627-636.

Ellegren, H., 2004. Microsatellites: simple sequences with complex evolution.. *Nature reviews genetics*, V(6), pp. 435-445.

Fluidigm, 2018. *SNP Genotyping: User guide*. San Francisco: Author.

González, E., 2003. Microsatélites: sus aplicaciones en la conservación de la biodiversidad. *Graellsia*, 59(2-3), pp. 377-388.

Infoagro, 2002. *Infoagro*. [En línea] Available at: [http://www.infoagro.com/frutas/frutas\\_tradicionales/melocoton.htm](http://www.infoagro.com/frutas/frutas_tradicionales/melocoton.htm) [Último acceso: 10 agosto 2018].

Kosambi, D., 1944. The estimation of map distances from recombination values.. *Annals of Eugenics*, 12(1), pp. 172-175.

Magi, A. y otros, 2014. H3M2: detection of runs homozygosity from whole - exome sequencing data.. *Bioinformatics*, 30(20), pp. 2852-2859.

Martínez-García, P. J. y otros, 2013. High density SNP mapping and QTL analysis for fruit quality characteristics in peach (*Prunus persica* L.). *Tree Genetics & Genomes*, Issue 9, pp. 19-36.

Meneses, C. y otros, 2016. A codominant diagnostic marker for the slow ripening trait in peach. *Molecular Breeding*, 36(77).

Meneses, C. y otros, 2016. A codominant diagnostic marker for the slow ripening trait in peach. *Molecular Breeding*, 36(77).

Núñez-Lillo, G. y otros, 2015. Identification of candidate genes associated with mealiness and maturity date in peach [*Prunus persica* (L.) Batsch] using QTL analysis and deep sequencing.. *Tree Genetics & Genomes*, Issue 11, p. 86.

Observatori de la fruita fresca, 2018. *Informe anual 2017*, Barcelona: Generalitat de Catalunya.

Osatinsky, R., 2007. ¿Que es la electroforesis capilar?. *Bioquímica y Patología Clínica*, 71(2), pp. 60-66.

Pérez, M., 2012. *Herramientas biotecnológicas aplicadas a la mejora genética de melocotón y nectarina*, Cartagena (España): Universidad Politécnica de Cartagena.

Picó Sirvent, M. B. & Esteras Gómez, C., 2011. *Marcadores moleculares basados en PCR: Marcadores SSR o STR.*, València: Universitat Politècnica de València.

Tautz, D., 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers.. *Nucleic Acids Research*, Issue 17, pp. 6463-6471.

Verde, I. y otros, 2017. The Peach v2.0 release: high-resolution linkage mapping and deep resequencing improve chromosome-scale assembly and contiguity. *BMC Genomics*, Issue 18, p. 225.